Essays in International Trade

Max Marczinek



A thesis submitted for the degree of

Doctor of Philosophy

University of Oxford

Department of Economics

Nuffield College

April 29, 2026

Contents

Abstract **iii**

List of Figures **vii**

List of Tables **xi**

1 Labour Scarcity and Productivity: Insights from the Last Nordic Plague **5**

1.1 Introduction 6

1.2 Context and Data 12

1.3 Trade 18

1.4 A Ricardian Model 27

1.5 Mechanism 32

1.6 Alternative Mechanisms 45

1.7 Counterfactual 52

1.8 Conclusion 53

2 Networks in Trade - Evidence from the Legacy of the Hanseatic League **55**

2.1 Introduction 56

2.2 Historical overview 60

2.3 Data 65

2.4 Hypotheses 69

2.5 Descriptives and bilateral trade data 72

2.6 Results 74

2.7 Probing Mechanisms 78

2.8 Conclusion 81

3 AI and Services-Led Growth: Evidence from Indian Job Adverts **87**

3.1 Introduction 88

3.2 Data . . . 94
3.3 Descriptives . . . 97
3.4 Empirical Strategy . . . 102
3.5 Main Results . . . 108
3.6 Robustness and Extensions . . . 116
3.7 Conclusion . . . 122

A Appendix for Chapter 1 **126**
A.A Data Appendix . . . 126
A.B History and Geography Appendix . . . 136
A.C Trade Appendix . . . 148
A.D Model Appendix . . . 177
A.E Mechanisms Appendix . . . 182
A.F Counterfactuals . . . 213

B Appendix for Chapter 2 **214**
B.A Data Appendix . . . 214
B.B Further results and robustness checks . . . 218
B.C Home town-level Estimation- Further Results and Robustness checks . . . 242

C Appendix for Chapter 3 **245**
C.A Data and Descriptives . . . 245
C.B Additional Results and Robustness . . . 260

Bibliography **293**

Abstract

The three chapters of this thesis study how trade and labour markets respond to shocks to factor proportions, trade costs, and technology. I show that such disruptions leave lasting imprints on the structures of comparative advantage and the labour market. The first chapter examines a natural experiment that substantially increased labour scarcity. Drawing on granular trade data, I show that the last outbreak of the Black Death in Northern Europe triggered capital deepening and productivity growth. This shifted the comparative advantage of affected regions towards capital-intensive sectors. I find that the effects outlast population recovery and provide additional evidence from production data. The second chapter studies the dissolution of the Hanseatic League, a merchant network that used to dominate trade in Northern Europe. We find that trade volumes converged to predictions from a gravity model. However, the network shaped the composition of trade for centuries. Long after the institution had formally disappeared, we show that trade on core routes remained concentrated in the hands of merchants whose families used to belong to the network. Together, these chapters establish that disruptions to factor markets and trade networks can produce lasting shifts in fundamentals. The third chapter brings this perspective to the present by documenting how the diffusion of artificial intelligence is affecting the Indian labour market. We find that AI adoption reduces labour demand and wage growth, and disproportionately so in high-skilled, non-routine, and analytical occupations. Across all three chapters, several unifying insights emerge: changes to production - whether driven by pandemics altering factor proportions, the institutional collapse of a merchant network, or innovative labour-saving technologies - affect trade, labour markets, and economic growth in deep and durable ways. This thesis expands our understanding of the mechanisms and persistence of these changes by drawing on both historical and contemporary settings.

Acknowledgements

This dissertation would not have been possible without the support, advice, and dedication of those around me. First and foremost, I thank my wife, who has been part of this journey from the very beginning. Through countless dinner-table conversations, she has likely become the world's second-leading expert on the intersection of early-modern Northern European plagues and multi-sector Ricardian models. I thank her for her love, support, and interest in my work, even when diverging into the history of draught animals or the estimation of scale elasticities.

I owe an enormous debt to my supervisors, who invested so much time and patience into me and my projects. I thank Ferdinand Rauch, who was my MPhil (and initial DPhil) supervisor and profoundly shifted my research interests in one conversation in the Old Quad of Brasenose College. I thank Steve Broadberry, whose big-picture insights and dedication to detail alike have been invaluable. Steve's work had fascinated me ever since the second term of my undergraduate degree; it has been even more inspiring to be under his supervision.

Banu Demir has shaped me as a researcher more than anyone else: when discussing the Hanseatic League chapter of this dissertation with her a few years ago, I realised that I, too, wanted to become a trade economist. It fills me with pride and happiness to say: mission accomplished!

Beyond my supervisors, I have benefited from a wider community at Nuffield College. I am grateful to my College supervisor, Steve Bond, who has supported me in manifold ways, and to my friend and co-author, Kristen, with whom I have started several exciting projects. Ruairí, Robin, Otso, Tanisha, and Victoria, you have made my time at Nuffield truly wonderful. I will always think back fondly to the first Covid lockdown — despite its sad backdrop, our little community was one of the greatest friend groups I have known.

I have also been fortunate in my collaborators. I thank Stephan Maurer, Alex Copestake, Ashley Pople, and Katherine Stapleton, who have been fantastic co-authors on these chapters. Beyond this dissertation, I am intellectually enormously indebted to Réka Juhász and Nathan Lane and thank them for sharing their passion for research. I hope to write many more papers

with you and the co-authors I have met at later stages of my dissertation.

I have spent long stretches of my DPhil abroad. In Berlin, I first and foremost thank Gabriel Ahlfeldt for hosting and welcoming me; he really created an academic home away from Oxford for me. I also thank Fan, Dimitria, Immanuel, and my other friends and colleagues at the department. In Stuttgart, I thank Nadja Dwenger and her chair for welcoming me so kindly. I thank the Department of Economic History at Uppsala for hosting me earlier in 2026. Special thanks go to the Forum for Research in Empirical International Trade and especially to Jon and Tibor, who have organised some of the best conferences I have ever attended.

Finally, there are so many more – family, friends, collaborators – whom I have not named here. I trust that you know who you are; I will always be grateful!

Declaration

Supervision This dissertation was supervised by Professor Banu Demir and by Professor Stephen Broadberry.

Word Count approx. 97,100 words

Statement of authorship I certify that this dissertation is my own work except where otherwise clearly stated.

Chapter 1 is solely my own work. Chapter 2 is joint work with Ferdinand Rauch and Stephan Maurer. An earlier version was submitted as part of the degree of MPhil in Economics at the University of Oxford, but has since been heavily revised. It is now published in the Journal of International Economics; I am including the latest working paper version. For this paper, I was responsible for the data cleaning and data analysis and earlier versions of the theoretical model. The conception, write-up, and later iterations on the model and empirical approach of the paper are joint work. Chapter 3 is joint work with Alexander Copestake, Ashley Pople, and Katherine Stapleton. On the empirical side, I have been responsible for improving causal identification, adding an event study and a granular task-level perspective, and for a wide range of robustness checks and addressing extensive sets of reviewer comments. The write-up of the paper and the conceptualisation and presentation of our findings are joint work. We are now finalising a re-submission to the Journal of Human Resources.

List of Figures

1.2.1 Plagued regions 14
1.2.2 Goods by sector 17
1.3.1 Capital-intensive exports expand relative to labour-intensive exports 22
1.3.2 Market shares expand 24
1.6.1 Scaling factor required to turn δ_{ikt} into pure Ricardian term 47
1.7.1 Counterfactual market shares without productivity channel 53

2.2.1 Red circles show Hansa cities according to Dollinger (1964), green circles show Kontore according to Hammel-Kiesow (2000); all other cities and towns in yellow. 63
2.3.1 This map shows the Øresund, where the Sound tolls were collected and all ships recorded, and the Great and Little Belts, the main alternatives for entering or leaving the Baltic Sea. 65
2.8.1 Hansa trade over time 83
2.8.2 Aggregate bilateral results 85
2.8.3 Hometown-level results 86

3.1.1 AI share of online job postings, by country 89
3.3.1 Trends in AI demand 99
3.3.2 Share of all AI posts by district, 2010-2019 100
3.3.3 Hiring profile of AI vs. non-AI vacancies 102
3.4.1 AI exposure by occupation group and by occupation wage offers 104
3.4.2 Impact of AI exposure on the AI share of establishments' posts 106
3.5.1 Impact of 1 unit higher establishment AI demand growth on verb usage by class and section 113
3.5.2 Impact of AI adoption on the wage offer distribution in non-AI posts 115
3.5.3 Impact of AI adoption on the non-AI wage offer distribution, holding occupational composition fixed 116

3.6.1 Event study for non-AI postings following AI adoption 119

A.B.1 Areas 138
A.B.2 Army marching routes, 1706-1714 140
A.B.3 Number of births, deaths, and marriages in Scania, 1647-1750 146
A.C.1 Number of ships passing the Sound per year 149
A.C.2 Alternative assignment as capital-intensive: Fact #1 150
A.C.3 Fact #1 using imputation estimator in Borusyak, Jaravel & Spiess (2021) . . . 151
A.C.4 Fact #1: Shift into capital-intensive exports by good 153
A.C.5 No shift into land-intensive exports after the plague 154
A.C.6 Capital-intensive exports expand relative to labour-intensive exports also in the long run 155
A.C.7 Robustness results for Fact #2: Export volumes and the extensive margin expand 156
A.C.8 Robustness results for Fact #2: Thirty pre periods and underlying value 157
A.C.9 Robustness results for Fact #2: NUTS 1-time and NUTS 2-time fixed effects . 158
A.C.10 Intensive margin expansion, volumes by sector 158
A.C.11 Lagged temperature and imputation estimator in Borusyak, Jaravel & Spiess (2021) 159
A.C.12 Intensive margin expansion, volumes by factor intensity 159
A.C.13 Fact #2: Volume expansion by good, capital-intensive sectors 160
A.C.14 Market shares and export volumes expand in the long run 161
A.C.15 Extensive margin expansion, by sector 162
A.C.16 Extensive margin expansion, by factor intensity 163
A.C.17 Fact #3: Extensive margin expansion by good, capital-intensive sectors 163
A.C.18 New exports at the extensive margin, by factor intensity 164
A.C.19 Fact #4: Increased probability of exporting a good for the first time, capital-intensive sectors 165
A.C.20 Plagued regions also expanded their imports 172
A.C.21 Map of Plagued NUTS 3 Regions 175
A.D.1 Simulated productivity growth and trade increase in observed vs. full sample . 181
A.E.1 Model-implied wage growth by mortality rate and labour share 185
A.E.2 Dynamics of population recovery after the plague: number of captains as proxy for population 185

A.E.3 Dynamics of capital accumulation after the plague: number of ships as proxy for the capital stock . . . 188
A.E.4 Imports of capital goods increase after the plague, suggesting capital accumulation 189
A.E.5 Agricultural production in Scania becomes more capital-intensive . . . 190
A.E.6 New exports at the extensive margin . . . 192
A.E.7 Scaling factor if labour shares assumption is violated . . . 193
A.E.8 The Plague's Productivity Effect by Sector, using Conley Standard Errors . . . 194
A.E.9 Total value of agricultural production per farm, controlling for farm size . . . 196
A.E.10 Share of Swedish Wharf Output by Plague Status . . . 196
A.E.11 Regions by serfdom status . . . 197
A.E.12 Urban-rural wage differential in Copenhagen . . . 198
A.E.13 Event studies testing for non-homotheticity . . . 212

B.A.1 Theoretical and observed distributions of distances to the Sound . . . 216
B.B.1 Ships from Hanseatic home towns did not differ in size from other ships . . . 219
B.B.2 Decade-interacted distance coefficients . . . 223
B.B.3 Simulation of bilateral trade: simulated Hansa effect recovered both in observed and in theoretical sample . . . 224
B.B.4 Aggregate bilateral results on tax value . . . 224
B.B.5 Aggregate bilateral results, OLS . . . 225
B.B.6 Aggregate bilateral results, only attendees of the last Hansetag and only active attendees of Hansetage . . . 226
B.B.7 Simulation of trilateral trade: a Placebo Hansa does not produce significant results 229
B.B.8 Home town-level results on tax value . . . 229
B.B.9 Home town-level results, OLS . . . 230
B.B.10 Home town-level results, excluding short distances . . . 231
B.B.11 Home town-level results, only attendees of the last Hansetag and only active attendees of Hansetage . . . 232
B.B.12 Home town-level results, excluding top 5 Hansa cities . . . 233
B.B.13 Home town-level results, only frequently used connections . . . 233
B.B.14 Home town-level results, controlling for time-interacted distances between either the origin and home town or destination and home town (whichever is shorter) . 234
B.B.15 Home town-level results, controlling for same country dummy . . . 236

B.B.16 Network results require Hanseatic captains and both Hanseatic origin and destination . . . 240
B.B.17 Home town-level results, distinguishing strict Hansa dummy from Hansa-Kontor dummy . . . 241
B.C.1 Home town-level results from multinational production and sectoral gravity models 244

C.A.1 Total posts by district, 2010-2019 . . . 246
C.A.2 Comparison of vacancy data with Prowess firm-level data and labor force surveys 249
C.A.3 Industry shares in vacancy data and labor force surveys over time . . . 251
C.A.4 Histogram of employment per firm (Prowess) and vacancies per firm (vacancies data) . . . 252
C.A.5 Histograms of key variables . . . 254
C.A.6 Trends in AI demand, number of posts . . . 256
C.A.7 AI posts by firm size . . . 257
C.A.8 Shares of posts by city . . . 257
C.A.9 Cities' shares of AI posts over time . . . 258
C.A.10 Top 20 roles demanding AI skills, 2010-2019 . . . 258
C.A.11 AI share of posts, by establishment age . . . 259
C.A.12 AI exposure across industries . . . 259
C.B.1 Impact of AI adoption on the wage offer distribution in all posts . . . 272
C.B.2 Non-AI vacancies following AI adoption at the district level . . . 277

List of Tables

1.3.1 Pre-plague balance checks 21
1.4.1 Spatial Separation & Capital Intensity 29
1.5.1 Impact of plague on sectoral productivity growth, by sector and factor intensity 41
1.5.2 Parish-level TFP growth after the plague 42

2.8.1 Key descriptives 84
2.8.2 Hanseatic captains continued to relatively dominate their old trade routes 85

3.2.1 Key descriptives of vacancy data, all jobs and AI jobs only 95
3.3.1 Verbs that are more (less) common in AI posts 101
3.5.1 Second stage: Impact of AI adoption on establishment non-AI vacancies 108
3.5.2 Second stage: Impact of AI adoption on establishment non-AI vacancy shares, by occupation group 110
3.5.3 Second stage: Impact of AI adoption on establishment routine and non-routine tasks 111
3.5.4 Second stage: Impact of AI adoption on establishment non-routine tasks, by occupation group 111
3.5.5 Second stage: Impact of AI adoption on establishment non-AI wages 114
3.6.1 Second stage: Impact of AI adoption on firm non-AI vacancies 120
3.6.2 Second stage: Impact of AI adoption on firm non-AI wages 121
3.6.3 AI adoption at the district level 121

A.A.1 Descriptive Statistics 127
A.A.2 Top 30 goods by value 130
A.A.3 Goods Classification 131
A.A.4 Currency conversion rates 132
A.A.5 Conversion of units of account 133
A.A.6 An example from the Soundtoll Data to show the different duties and sub-totals 134
A.B.1 Areas and NUTS correspondence 137

A.B.2 List of plagued cities and mortality estimates 141
A.B.3 List of besieged cities 142
A.B.4 Determinants of urban mortality 142
A.B.5 Impact of plague on city population 144
A.B.6 Parish-level population recovery after the plague 145
A.B.7 Impact of plague on number of births in 59 Scania parishes 146
A.C.1 Naval raids and blockades affecting Baltic ports, 1709-1721 148
A.C.2 Proximity to Christiansø is associated with significantly lower exports 149
A.C.3 Plague increases capital-intensive exports more than labour-intensive exports . 151
A.C.4 Shift into capital-intensive exports by army proximity and siege 152
A.C.5 Plague increases capital-intensive exports more than labour-intensive exports in both sectors 154
A.C.6 Plague increases exports at the intensive and extensive margins 157
A.C.7 Impact of plague on number of goods never exported before, by factor intensity 166
A.C.8 Impact of plague on volume of goods never exported before, by factor intensity 166
A.C.9 Gravity analysis of plague shock on bilateral trade 167
A.C.10 Impact of plague on probability to export 168
A.C.11 Heterogeneity of trade findings by pre-plague export levels 169
A.C.12 Travel times are unaffected by the plague 170
A.C.13 Heterogeneity of trade findings by siege 171
A.C.14 Export expansion and shift into capital-intensive goods in post-plague Hamburg 172
A.C.15 List of plagued cities and regions across all of Europe 174
A.C.16 Impact of Plague on Trade, Accounting for Outbreaks in the Hinterland 176
A.E.1 Grain price-inferred interest rates in Danzig do not increase 186
A.E.2 Examples of capital deepening by sector 187
A.E.3 Impact of plague on proxied capital stock 188
A.E.4 Farm production and distance to plagued cities 190
A.E.5 Parish-level shift into higher capital intensity within agriculture 191
A.E.6 Farm production, distance to plagued cities, and soil quality 191
A.E.7 Impact of plague on sectoral productivity growth in the long run 195
A.E.8 Impact of plague on shift into capital-intensive exports by sector & second serfdom 199
A.E.9 Impact of plague on sectoral productivity growth, by second serfdom 200

A.E.10 Pre-war financial development (proxied by export-weighted APP) predicts larger treatment effects . . . 202
A.E.11 Impact of plague on sectoral productivity growth, by financial development . . 202
A.E.12 The plague leads to a shift into physical capital intensity, not into human or social capital intensity (proxied by skill and contract intensity) . . . 204
A.E.13 Product-level prices by sector and year in Danzig and Amsterdam . . . 206
A.E.14 Good-level logarithmic import prices . . . 208
A.E.15 Good-level logarithmic export prices . . . 210

B.B.1 List of Hansa and Kontor cities appearing in the trade data . . . 218
B.B.2 Hansa and city population . . . 220
B.B.3 Hanseatic Cities Imported More and More Luxury Goods . . . 221
B.B.4 Regression results depicted in Figure 2.8.2. Point estimates for β_t and λ_t in equation 2.5.1 and regression summary statistics . . . 222
B.B.5 Overview of aggregate bilateral results without time dimension . . . 227
B.B.6 Regression results depicted in Figure 2.8.3a. Point estimates for β_t and δ_t in equation 2.6.1 and regression summary statistics . . . 228
B.B.7 Overview of home town level results without time dimension . . . 235
B.B.8 Hansa routes connect geographically more similar cities . . . 237
B.B.9 Overview of home town level results without time dimension, adding origin x home town and destination x home town fixed effects . . . 237
B.B.10 No evidence that Hanseatic captains used more standardised units . . . 238
B.B.11 Hanseats traded a larger variety of goods on Hanseatic ships . . . 239

C.A.1 Number of observations by data source . . . 246
C.A.2 Vacancies data at different stages of data restriction . . . 253
C.A.3 Difference (2017-19 over 2010-12) descriptives of vacancy data in restricted sample, all jobs and AI jobs only . . . 253
C.A.4 Wages in AI vs. non-AI roles . . . 255
C.A.5 Verbs in AI posts and example nouns . . . 257
C.B.1 Investigating correlates of shares . . . 261
C.B.2 Examining pre-trends for the instrument . . . 262
C.B.3 Second stage: Impact of AI adoption on establishment non-AI vacancies. Adão, Kolesár, and Morales (2019) standard errors. . . . 263

C.B.4 First stage: Impact of AI exposure on establishment AI adoption 264
C.B.5 Second stage: Impact of AI adoption on establishment non-AI vacancies, excluding ten largest firms 265
C.B.6 Second stage: Impact of AI adoption on establishment non-AI wages, excluding ten largest firms 265
C.B.7 Second stage: Impact of AI adoption on establishment non-AI vacancies, excluding the three establishments accounting for most of the 2016-17 growth in AI vacancies 266
C.B.8 Second stage: Impact of AI adoption on establishment non-AI wages, excluding the three establishments accounting for most of the 2016-17 growth in AI vacancies 266
C.B.9 Second stage: Impact of AI adoption on establishment non-AI vacancies, excluding the 5 establishments contributing two thirds of AI job posts 267
C.B.10 Second stage: Impact of AI adoption on establishment non-AI wages, excluding the 5 establishments contributing two thirds of AI job posts 267
C.B.11 Second stage: Impact of AI adoption on establishment non-AI vacancies, excluding top decile of firms by adoption 268
C.B.12 Second stage: Impact of AI adoption on establishment non-AI wages, excluding top decile of firms by adoption 268
C.B.13 Second stage: Impact of AI adoption on establishment non-AI vacancies, excluding finance sector 269
C.B.14 Second stage: Impact of AI adoption on establishment non-AI wages, excluding finance sector 269
C.B.15 Second stage: Impact of AI adoption on establishment non-AI vacancy shares, by granular occupation group 270
C.B.16 Second stage: Impact of AI adoption on establishment abstract and routine tasks 270
C.B.17 Second stage: Impact of AI adoption on establishment routine and non-routine tasks, holding occupational composition fixed 271
C.B.18 Second stage: Impact of AI adoption on establishment abstract tasks, by occupation group 271
C.B.19 Second stage: Impact of AI adoption on establishment non-AI wages, by occupation group 272
C.B.20 Second stage: Impact of AI adoption dummy on establishment non-AI wages, controlling for job profiles 273

C.B.21 Second stage: Impact of AI adoption on establishment non-AI vacancies, controlling for baseline share of software engineers 274
C.B.22 Second stage: Impact of AI adoption on establishment non-AI wages, controlling for baseline share of software engineers 275
C.B.23 Second stage: Impact of AI adoption on establishment non-AI vacancies, controlling for baseline share of sales & admin vacancies 275
C.B.24 Second stage: Impact of AI adoption on establishment non-AI wages, controlling for baseline share of sales & admin vacancies 276
C.B.25 Second stage: Impact of AI adoption on establishment non-AI vacancies, 2013-15 to 2017-19 276
C.B.26 Second stage: Impact of AI adoption on establishment non-AI wages, 2013-15 to 2017-19 277
C.B.27 Vacancies results for 'incumbents' and 'startups' 278
C.B.28 Wage results for 'incumbents' and 'startups' 279
C.B.29 Second stage: Impact of AI adoption on establishment non-AI vacancies, intensive margin of AI demand 280
C.B.30 Second stage: Impact of AI dummy on establishment non-AI wages, intensive margin of AI demand 280
C.B.31 Second stage: Impact of AI adoption on establishment non-AI vacancies, IHS of AI vacancies 281
C.B.32 Second stage: Impact of AI adoption on establishment non-AI vacancies, IHS of AI vacancies 281
C.B.33 Second stage: Impact of AI adoption on establishment non-AI vacancies and wages, including firm fixed effects 282
C.B.34 Second stage: Impact of AI adoption on establishment non-AI vacancies, excluding last year of data 282
C.B.35 Second stage: Impact of AI adoption on establishment non-AI wages, excluding last year of data 283
C.B.36 Second stage: Impact of AI adoption on establishment non-AI vacancies, weighted (top 5% winsorized) 283
C.B.37 Second stage: Impact of AI adoption on establishment non-AI wages, weighted (top 5% winsorized) 284

C.B.38 Second stage: Impact of AI adoption on establishment non-AI vacancies, weighting vacancies to match labour survey occupation shares 284
C.B.39 Second stage: Impact of AI adoption on establishment non-AI wages, weighting vacancies to match labour survey occupation shares 285
C.B.40 Second stage: Impact of AI adoption on establishment non-AI vacancies, keeping only occupations with vacancy shares similar to labour survey occupation shares 285
C.B.41 Second stage: Impact of AI adoption on establishment non-AI wages, keeping only occupations with vacancy shares similar to labour survey occupation shares 286
C.B.42 Second stage: Impact of AI adoption on establishment non-AI vacancies, two-step estimation method 286
C.B.43 Second stage: Impact of AI adoption on establishment non-AI wages, two-step estimation method 287
C.B.44 Second stage: Impact of AI adoption on establishment non-AI mean wages . . 287
C.B.45 OLS: Regression of AI adoption on establishment non-AI vacancies 288
C.B.46 Probit regression on establishment AI adoption 289
C.B.47 Second stage: Impact of AI adoption on establishment non-AI vacancies – Felten et al. exposure measure 291
C.B.48 Second stage: Impact of AI adoption on establishment non-AI vacancy shares, by occupation group – Felten et al. exposure measure 291
C.B.49 Second stage: Impact of AI adoption on establishment non-AI wages – Felten et al. exposure measure 292
C.B.50 First stage: Impact of AI exposure on establishment AI adoption – alternative exposure measures 292

Introduction

This dissertation consists of three chapters in international trade. What ties them together is a concern with how trade and labour markets respond to large structural shocks, and with how long the resulting changes persist. A pandemic that removes a substantial fraction of the workforce, the dissolution of a centuries-old merchant network, and the spread of artificial intelligence in a developing country are very different events. Yet each, I argue, reshapes production and trade in persistent ways which this dissertation seeks to elucidate.

The first two chapters share a data source as both rely on the Sound Toll Registers, a remarkable record of cargoes passing through the Øresund between Denmark and Sweden. These registers were kept by the Danish crown from 1497 to 1857. For every voyage, the registers list the captain, the home port, the origin and destination, and the cargo along with the toll paid. The result is a panel of trade flows of extraordinary geographical and temporal resolution – closer in detail to modern customs data than to anything else available for the pre-industrial world. Both chapters exploit this resolution to study the economic effects of shocks at a finer spatial scale than was previously feasible: a plague outbreak that struck some Baltic ports but spared their neighbours, and a network whose membership was defined at the city level. The third chapter shifts setting and method but takes up a related question: when a novel disruptive technology arrives, who benefits, who is displaced, and how is the shock transmitted across the labour market?

I begin in chapter 1 with the question of whether labour scarcity can be a spur to productivity growth. Settling the question empirically is hard, because labour-supply shocks rarely arrive in isolation and are historically typically not documented in sufficient detail. The 1708–1712 Great Northern War plague offers an unusually clean setting. It struck some Baltic ports with devastating mortality and bypassed others, and it occurred decades before the diffusion of any of the technologies associated with the Industrial Revolution. By comparing trade patterns across plagued and unplagued ports of similar size and geography, it is possible to isolate the effects of labour scarcity itself.

What I find is that plagued regions did not lose competitiveness; they gained it. Their share of capital-intensive exports rose by about ten percentage points relative to a pre-plague baseline of 6.6%, and their shares of destination markets grew by roughly one percentage point against a baseline of 0.9%. The shift towards capital-intensive goods is consistent with a Rybczynski-style reallocation: when labour becomes scarce, the relative abundance of capital rises, and capital-intensive sectors expand. But the broader market-share gains – which appear across sectors and factor intensities, and along both the intensive and extensive margins – cannot be generated by factor proportions alone. To make sense of both findings, I build a Ricardian model with immobile labour, in which capital deepening raises productivity and does so most strongly in capital-intensive sectors. The model's predictions match the data, and I confirm the productivity channel directly using Swedish farm-level production records. The patterns I document persist for almost a century, even though populations had returned to their pre-plague trends within forty years. The plague's effects on comparative advantage were not undone by demographic recovery.

Chapter 2 turns to a different sort of disruption: the slow institutional decline of a centuries-old commercial network. The Hanseatic League linked merchants from Bruges to Novgorod for roughly four hundred years, and at its height it dominated trade in the Baltic and North Seas. By 1669, however, it had been formally dissolved. With the Sound Toll data spanning nearly two centuries after this dissolution, we can ask what happened to trade as the League disappeared. The answer comes in two parts. First, trade volumes on former Hanseatic routes converged to what a standard gravity model would predict. There is no permanent volume premium left over from the network. Second, trade composition – specifically, who carried trade between which cities – did not converge. We track captains by family name and home town and find that on the League's traditional core routes, trade remained concentrated in the hands of merchants from Hanseatic families generations after the institution itself had ceased to exist. The formal network dissolved; the relational network it had built outlived it by a long margin.

Chapter 3 leaves the early-modern period for the present and asks: when a new and disruptive production technology arrives, what happens to the workers who used to do the work? We study artificial intelligence in the Indian services sector, which has been the engine of much of India's recent growth. The data are fifteen million online vacancy posts from India's largest jobs board, covering the period from 2010 onwards. From the text of the postings we identify which jobs require AI-related skills, which establishments are hiring for them, and, most granularly, which exact tasks are demanded. The descriptive picture is striking. AI-skilled vacancies pay

substantially more, demand more education, and are concentrated in a small number of firms and cities. Demand for these skills has grown rapidly since around 2016, particularly in IT, finance, and professional services.

The descriptive picture, on its own, does not tell us whether AI is creating jobs, destroying them, or reshaping them. To get at that question, we construct a shift-share instrument. We find that in firms that take up AI, the growth of non-AI vacancies slows, total hiring slows, and wage offers fall across the distribution. The effects are largest in occupations that are high-skilled, non-routine, and analytical. These are precisely the categories where AI's recent advances have been most visible. Whether new firms and new tasks elsewhere in the economy will absorb the displaced workers is a question the data we use cannot fully answer. But the picture is clear enough to raise real concerns about whether the services-led growth path will continue to generate jobs at the rate India's demographic situation requires.

Contributions. The dissertation contributes to several literatures across trade, growth, economic history, and labour economics.

Chapter 1 contributes to a long-standing debate about the relationship between labour scarcity and productivity. The Black Death has been studied extensively, but most of that work has focused on wages, urbanisation, or aggregate output, and has used data that are too coarse to identify sectoral reallocation. By exploiting port-level cargo records, I provide direct evidence on how plagued regions' export composition changed, and on how those changes endured. I show that comparative advantage structures remained altered even when population levels had recovered. This is a novel contribution in that I find that a transient shock to factor proportions can lead to a persistent change in trade patterns and productivities. Further, by combining granular trade data, a Ricardian framework, and Swedish farm- and parish-level production data, I contribute to an improved understanding of the mechanisms behind economic recovery from large shocks.

Chapter 2 makes two contributions. First, it constructs what is, to our knowledge, the most granular panel of early-modern trade flows yet used in the literature, covering nearly two centuries of voyages between several hundred port pairs. Second, it separates the volume effects of network dissolution from the composition effects, a distinction that previous work on merchant networks has largely been unable to make. The paper is able to do this thanks to the exceptional granularity of the Soundtoll data, which allows us to identify individual captains. Thus, we extend existing gravity studies by introducing the margin of individual merchants. The finding that informal ties survive the formal institutions that produced them has implications beyond

the Hanseatic case, including for current debates about the resilience of trading relationships when geopolitical alignments shift.

Chapter 3 also makes two contributions. First, it assembles a large new dataset on Indian online vacancies that includes wages, education, and experience requirements, allowing for a more detailed picture of AI demand in a developing economy than has previously been available. Second, it estimates the causal effect of establishment-level AI adoption on hiring and wages, which is, to our knowledge, the first such estimate for a developing country. Most of the existing literature is concentrated on the United States and the United Kingdom; our results both extend the geography and add further granularity by studying how AI adoption affects task demand.

Structure. The dissertation is organised as follows. Chapters 1 to 3 present the essays on the Last Nordic Plague, the Hanseatic League, and AI in India. The final chapter concludes. Each chapter is self-contained; chapter-specific appendices follow at the end of the dissertation. This is where the introduction goes.

Chapter 1

Labour Scarcity and Productivity: Insights from the Last Nordic Plague

Max Marczinek[1]

Abstract. I study the causal impact of labour scarcity on productivity growth, a key driver of long-run economic growth. Exploiting a natural experiment from a 1710s plague outbreak in Northern Europe, I show that plagued regions shift into capital-intensive exports and see an export expansion. Using rich port-level trade data and a Ricardian model, I trace this shift to productivity growth driven by capital deepening. While population levels recover within four decades, the productivity and trade effects persist for almost a century, suggesting long-run changes in comparative advantage. My findings imply that labour scarcity can induce productivity-enhancing reallocation.

JEL Classification Codes: F16, O47

Keywords: Labour Scarcity, Productivity, Growth

[1]I thank the Soundtoll team and Jürg Luterbacher and Ulrich Pfister for sharing their data. This research has been supported by a grant from the Department of Economics at the University of Oxford.

1.1 Introduction

What determines economies' long-run growth rate? One of the main drivers of economic growth are sustained increases in total factor productivity. A long-standing hypothesis holds that labour scarcity, by raising wages and altering relative factor prices, may spur productivity-enhancing innovation and reallocation (Voth et al. (2025), Habakkuk (1962)). However, the opposing view holds that labour scarcity hampers growth as firms are unable to scale production and competitiveness is harmed by higher wages (OECD (2024)). This debate has taken on an increased importance as advanced economies face both low productivity growth and a shrinking workforce.

In this paper, I study a context where labour scarcity leads to productivity growth. I combine granular geographical variation in labour scarcity from a natural experiment, the last Nordic plague, with novel port-level trade data. This allows me to document how labour scarcity affects the composition and volume of trade. I find that plagued regions shift their exports into capital-intensive goods and conquer larger market shares abroad. While the prior reflects changes in factor proportions, I argue that a Ricardian productivity component is required to explain the latter. I thus rationalise these findings in a Ricardian model with immobile labour and link capital deepening to productivity growth. Finally, I show that these plague-induced comparative advantage changes are still in place half a century after population recovery.

The historical setting of this natural experiment is uniquely suited to study this question. I exploit the 1708-1712 Great Northern War plague outbreak as a plausibly exogenous shock that generated labour scarcity. While earlier plagues affected all of Europe, this later outbreak left many other Northern European regions unaffected. These nearby and similar regions form an ideal comparison group. Further, I study a period decades before the Industrial Revolution, which allows me to abstract from the diffusion and adoption of novel technologies.

Exceptionally detailed trade data allow me to track plagued regions' trade patterns for decades while rigorously incorporating the gravity structure of trade. Two features make the Danish toll records used in this paper the ideal source. First, origins and destinations were recorded at the level of ports rather than countries. I therefore observe trade at the same granular spatial resolution as the treatment. Second, cargoes were disaggregated into goods, which I assign into labour- and capital-intensive sectors. I can thus differentiate between goods of different factor intensities. In sum, these data permit an unprecedentedly detailed analysis of trade following this labour supply shock.

I establish two novel empirical findings. First, I find that plagued regions' share of capital-

intensive exports rises by 10pp, compared to a pre-plague mean of 6.6%. New capital-intensive export goods contribute significantly to this finding. As capital became relatively more abundant, the shift into capital-intensive exports is reminiscent of the Rybczynski theorem, which states that a relative rise in a factor's endowment leads to an expansion in sectors using that factor intensively. However, I argue that a factor proportions framework cannot explain this paper's findings.

Second, plagued regions expand their market shares abroad. Their destination market shares rise by 1pp, compared to a pre-plague mean of 0.9%. This finding is surprising as labour scarcity should increase wages and reduce regions' competitiveness. Instead, I find evidence for a pattern of forced experimentation and innovation as plagued regions also add novel goods to their export baskets. While the export expansion is strongest in capital-intensive sectors, it holds across sectors and factor intensities, along both the intensive and extensive margins. This suggests an important role for productivity growth in the post-plague market share expansion. To eliminate alternative explanations, I show a battery of robustness checks.

Finally, these trade results display path dependence beyond population recovery. By 1750, about four decades after the plague, populations had returned to their pre-plague trends. However, trade and comparative advantage patterns reflect past labour scarcity for almost a century. Thus, the plague's effects on trade patterns outlive population recovery by half a century. This suggests that adjustments to labour scarcity produce persistent changes in comparative advantages.

Next, I build a Ricardian model following Eaton & Kortum (2002) to rationalise my reduced-form findings. The model's Ricardian nature is motivated by the fact that exports expand across sectors and factor intensities, which a pure Heckscher-Ohlin model cannot explain. I also find direct evidence of post-plague TFP growth in production data. The proposed Ricardian model with immobile labour can reproduce both the shift into capital-intensive exports and the observed export expansion. As labour is immobile across regions, Ricardian forces are married to factor-content forces. The proposed model features two competing channels. First, a Ricardian channel that permits Hicks-neutral changes in productivity (Davis & Weinstein (2001)). I follow Costinot et al. (2011) in modelling Ricardian forces within sectors. Second, the model nests non-homothetic preferences as an alternative demand-side channel (Fieler (2011)). I use this model to quantify the relative importance of these channels.

I argue that productivity growth is best-suited to explain my findings. I propose a three-step mechanism that matches my empirical results. First, labour scarcity leads to higher wages. Second, producers employ relatively more capital and exports shift into capital-intensive goods.

Third, sectoral productivity growth accelerates as the capital-to-labour ratio rises. I argue that productivity growth as a Ricardian component is necessary to explain the export expansion across sectors and factor intensities. Thus, I suggest that capital deepening increases productivity, and I allow the elasticity between capital deepening and productivity to vary by sector. This reduced-form dependence of productivities on factor proportions allows the model to match the empirical findings.

In support of this mechanism, I employ granular agricultural data from Sweden to show capital deepening *within* agriculture - more horses per capita were used in labour-scarce parishes after the plague - and *across* sectors - agricultural production in labour-intensive arable farming drops and shifts into capital-intensive animal husbandry. Drawing on Swedish parish-level agricultural data, I additionally estimate TFP before and after the plague, and find that TFP growth is significantly larger in plagued parishes. I also construct a proxy for the depths of ports' capital markets and show that financially more developed plagued ports undergo stronger capital deepening. Further, I show that plagued regions add significantly more novel goods, not exported by them before, to their export baskets. This finding is significantly stronger in capital-intensive sectors, reflecting that investments in new production lines are associated with gains in productivity and competitiveness, and that this association is stronger in capital-intensive sectors.

While I treat productivities as fundamentals of the model, there are several micro-foundations for capital deepening increasing productivity. Leaning on Bustos (2011), I show that technology choice driven by factor prices produces the suggested relationship. As wages rise, more producers switch to capital-intensive technology which confers a productivity boost on them yet requires higher fixed costs. This increases both the capital-to-labour ratio and sectoral productivity. I show that in this framework sectoral productivity increases in capital deepening. Other possible micro-foundations include learning by doing (Krugman (1987)), cost discovery (Hausmann & Rodrik (2003)), and learning by exporting (Loecker (2013)). All these micro-foundations can produce an ascending relationship between productivity and capital deepening. From the model's gravity equation, I recover productivity growth as fundamentals (Costinot et al. (2011)) and project these recovered values onto changes in factor proportions, where institutions and geography are absorbed by fixed effects (Chor (2010)).

Using the model, I first show that wages rise as labour scarcity bites and validate this prediction with historical evidence. The shift into capital-intensive goods, the second step, is attributed to three channels by the model: differences in labour intensity between sectors; relative price effects passed on to destination markets; and sectors' different elasticities of productivity

with respect to the capital-to-labour ratio. While all sectors face the same regional factor proportions, they may differ in their labour shares and in their productivities' responsiveness to capital. In this historical setting, the latter mainly reflects the availability of productivity-enhancing technologies requiring capital. Such technologies were very limited in labour-intensive sectors (Atack et al. (2019), Gallardo & Sauer (2018), Coleman (1956)). Additional capital relative to labour should produce larger productivity increases in capital-intensive sectors, where such technologies were more readily available; thus, their productivities' elasticity with respect to capital deepening should be higher.

In support of the third step, I show that productivity growth accelerates after the plague. Capital-intensive agriculture and manufacturing see faster productivity growth than their labour-intensive counterparts. These productivity changes, too, last for almost a century and illustrate path dependence in comparative advantages. I thus argue that labour scarcity leads to productivity growth. However, there is a caveat: institutions shielding producers from these changes stand in the way of the necessary adjustments. I consider one such institution, serfdom, and show that my findings do not apply to regions with serfdom. Labour scarcity was most severe in cities, but serfs could not move there, keeping the hinterland artificially abundant in labour. The lack of productivity effects under serfdom suggests that producers did not face the right set of incentives.

I find no support for a competing pure Heckscher-Ohlin explanation and show that Ricardian forces account for more than 90% of sectoral differences on the supply side in a hybrid framework. Similarly, I find no evidence for a demand-side channel of non-homotheticity, which is frequently discussed in the plague literature (Voigtländer & Voth (2012)). My model nests non-homothetic preferences (Fieler (2011)), under which demand may shift into income-elastic goods as wages rise. However, I find no evidence for this in a direct test and in implied price effects. The absence of changes in import composition speaks against sector-specific embodied technology linked to increasing imports; it also provides another argument in favour of Ricardian and against pure Heckscher-Ohlin forces. Similarly, I argue against venting out and market power as explanations. Further, I find no evidence for human or social capital deepening after the plague: even when controlling for goods' skill and contract intensities, the only shift in exports after the plague remains into physical capital intensity. However, I acknowledge that factor-biased technological progress and a hybrid model containing both factor content and Ricardian forces could rationalise my findings. Such models would nonetheless require productivity growth after the plague. I thus conclude that the mechanism I propose is well-suited to explain my empirical findings. Finally, I establish that in the counterfactual absence of a productivity channel exports would have

contracted. The productivity growth necessary to bridge the counterfactual export contraction and the observed export expansion closely matches the estimated TFP growth figures. Therefore, I propose that productivity growth is necessary to reconcile theory with empirics.

This work builds on several strands of literature. First, I provide novel evidence that labour scarcity leads to productivity growth. I use an exogenous mortality shock, instead of migration or fertility that present endogeneity challenges, to establish these findings. I also go beyond existing work to link labour scarcity not only to the adoption of machines and innovation (as in Voth et al. (2025), Franck (2024), and Andersson et al. (2020)) but also to productivity growth. This is the ultimate outcome of interest in this debate to which my paper makes a crucial contribution.

Second, this paper makes a key contribution to the literature on path dependence in the spatial economy. My paper demonstrates that temporary labour scarcity is reflected in comparative advantages for almost a century. Thus, I argue that a shock to factor proportions explains persistent changes in comparative advantages. I bring Ricardian theory to unprecentedly granular trade data, which also allows me to go beyond existing work (Lane (2025), Juhász (2018)) in establishing path dependence in comparative advantages in a general equilibrium framework. Further, the historical setting of this natural experiment allows me to avoid the pitfalls of technology diffusion and policy endogeneity.

Third, I provide granular evidence and plausible micro-foundations for the mechanisms behind regions' recovery after shocks (Jedwab et al. (2024), Voigtländer & Voth (2012), Davis & Weinstein (2002)). I show that this recovery entailed active adjustments reflecting economic decisions rather than mechanical forces.

This paper proceeds as follows. The following section summarises the related literature. Section 1.2 introduces the historical context and the data. Section 1.3 holds the trade results and section 1.4 the model. In section 1.5, I discuss the mechanisms behind my results. Section 1.7 shows a counterfactual analysis of trade in the absence of a productivity channel and section 1.8 concludes.

Related Literature

Labour Scarcity & Productivity

The view that labour scarcity promotes productivity growth is well-established in the literature. Habakkuk's seminal work postulates that labour scarcity in the United States led to labour-saving technological progress (Habakkuk (1962)). Acemoglu (2010) generalises the relationship between factor proportions and the factor bias of technological progress. He shows that labour scarcity

encourages technological progress if the latter reduces the marginal product of labour. Allen (2009*a*) considers Britain a high wage, cheap energy economy and argues that British innovations resulted from these factor proportions.[2]

Recent work has added causal identification to this debate. Voth et al. (2025) use variation in warships' coastal access to identify the positive effect of labour scarcity on the adoption of labour-saving machines. Franck (2024) exploits temperature variation to analyse how labour scarcity led to technology adoption and innovation. Andersson et al. (2020) use an IV approach to identify the causal effects of emigration on technological change, linking higher capital intensity to innovation.

In current-day debates, labour scarcity is considered an obstacle to economic growth (OECD (2024), Caldara et al. (2024)). An emerging literature shows that firms respond to labour scarcity by competing for workers, adjusting their input mix, and innovating in management techniques (Börschlein et al. (2024), Groiss & Sondermann (2023)).

My paper's key empirical contribution is to test existing theories. In particular, my paper forms the first test of the Habakkuk thesis in an open economy setting. Further, the commodity- and port-level trade data set is an empirical contribution in itself. On the theoretical side, I link labour scarcity and capital deepening-driven productivity growth in a general equilibrium framework. This addresses the criticism of Temin (1966) towards Habakkuk (1962) regarding factor proportions in general equilibrium.

Path Dependence in Comparative Advantage

An established literature proposes micro-foundations for path dependence in comparative advantage. Krugman (1987) shows how learning-by-doing feeds into lasting productivity changes. A related explanation lies in learning by exporting. Clerides et al. (1998) find no general evidence that firms become more productive after entering export markets. Loecker (2013) argues to overcome previous identification challenges and finds that entering export markets leads to productivity growth. More generally, Allen & Donaldson (2022) build a model to explain persistent effects of temporary shocks.

A growing body of literature tests this empirically. Juhász (2018) interprets the British blockade of France as temporary trade protection. She finds that more protected regions increased their spinning capacity and had higher value added per capita in industry for several decades. Lane (2025) studies industrial policies in South Korea and argues that they successfully changed

[2]This assertion is subject to criticism by Humphries & Schneider (2020), Humphries & Weisdorf (2019), and Stephenson (2018).

comparative advantages.[3]

I contribute to this literature by showing that an exogenous increase in capital intensities is followed by a lasting shift in comparative advantages. Juhász, Lane & Rodrik (2024) argue that assessing policies is complicated by identification challenges. Instead, I study an exogenous shock, which largely exempts my findings from this criticism.

Economic Effects of Plagues

My paper also relates to the literature on plagues. Population recovery after the Black Death took about two centuries (Jedwab et al. (2024)). Wages, however, did not recover in all parts of Europe (Jedwab et al. (2022), Alvarez-Nogal et al. (2020)).[4] Living standards rose: Broadberry et al. (2014) document that the caloric composition in England shifted towards dairy and meat. Gelman (1982) finds that English agriculture shifted from arable into pastoral farming. Bosshart & Dittmar (2025) and Dittmar & Meisenzahl (2019) stress the political economy of plague outbreaks. Voigtländer & Voth (2012) argue that sustained wage increases were maintained by directly creating downward pressure on urban populations. Jedwab et al. (2022) document the increased bargaining power of workers beyond higher wages.

My paper adds to this literature by focusing on granular trade in a general equilibrium framework and establishing novel empirical facts. My findings reject the predictions of Malthusian models (Campbell (2010)) and of models of non-homothetic demand (Voigtländer & Voth (2012)). Instead, I find productivity growth effects that have been implicit to findings in existing work. A theoretical contribution lies in my model's ability to explicitly parse out non-homotheticity (Fieler (2011)). This paper thus helps to establish a firmer foundation of our understanding of the economic effects of plagues.

1.2 Context and Data

This paper studies the last plague outbreak in Northern Europe, which coincided with large-scale warfare. The war that broke out in 1700 between Sweden, Russia, and their allies affected most countries on the Baltic Sea. Early campaigns led the Swedish army into the Baltics and later deep into the Polish-Lithuanian Commonwealth. As the war turned in Russia's favour, Swedish troops receded to Northern Germany, and Russian troops marched far into Sweden. After the

[3] More broadly on the time path of changes in trade patterns, Hanson et al. (2018) stress the empirical relevance of dynamic comparative advantage.

[4] In the short-run, Jedwab et al. (2022) find that wages actually dropped. They explain this through Smithian growth going into reverse.

decisive Swedish defeat in Poltava in 1709, the Danish navy blocked ships between Sweden and its Northern German possessions, and the Russian navy raided the Baltic coast. As a result, ship passages at the Sound collapsed. This so-called Great Northern War ended in 1721 and marked the end of Sweden as a great power. Swedish territories in the Baltic were lost to Russia, and about half of Swedish Pomerania to Prussia.

The war was accompanied by a severe outbreak of the plague. The plague mostly followed army routes, reaching East Prussia in 1708, most of the Baltic Sea by 1711 and Hamburg by 1712. At this point, the Swedish army was on her way back from what is today Ukraine and sought refuge and fresh supplies in Swedish Pomerania. Appendix Figure A.B.2 shows digitised army marching routes based on Spruner & Menke (1880) and Barraclough (1997) to illustrate how armies spread the plague. This historical evidence speaks towards war, not trade, spreading the plague. Figure 1.2.1 shows plagued regions in my sample.[5] 22 regions contain cities that were besieged during the war, while seven regions were both plagued and besieged.[6] Most plagued cities caught the plague from their hinterlands through which armies marched.

1.2.1 Plague & Mortality Data

Appendix Table A.B.2 presents the plague data I collected. Virtually all plagued regions should be accounted for: the literature covers all areas around the Baltic, and it is reasonable to assume that a plague incident will have been recorded. Even Bremen's mortality rate of 0.7% is recorded. I restrict myself to coding plagued regions. While Raster (2023) presents hand-collected rural mortality data for Northern Estonia, this remarkable effort can hardly be extended to the entire region. For towns covered in both the trade data and in Raster (2023), the plague status overlaps precisely.

Mortality estimates are available for half of plagued cities, with the median mortality rate at 36%. For the remaining plagued cities, I predicted mortality with the timing of the plague, geographical controls, and the proximity to army routes. The proximity to armies, displayed in Appendix Figure A.B.2, is motivated by the finding that armies spread the plague. Appendix Table A.B.4 shows the regression results.

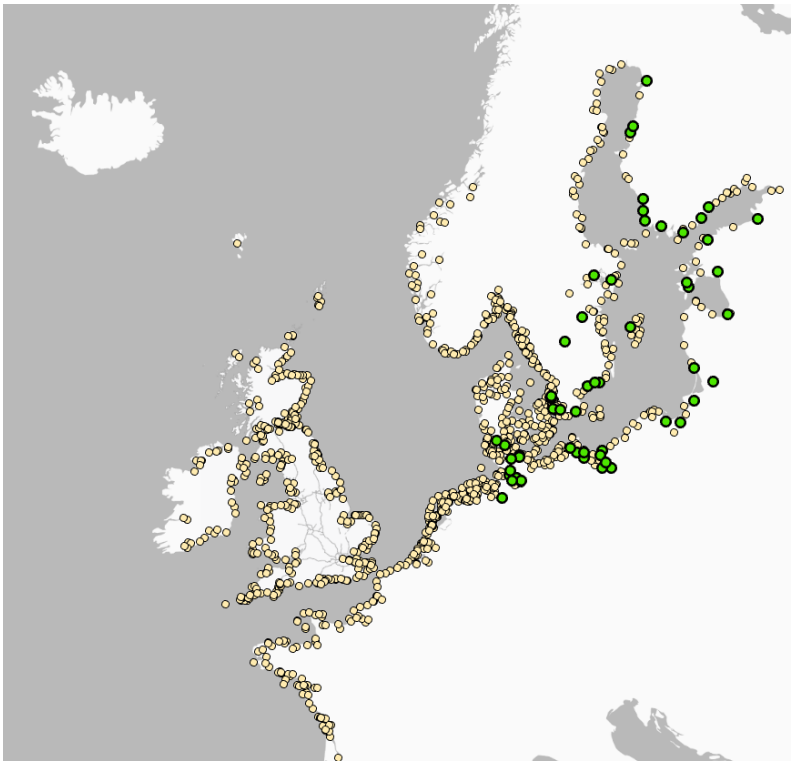
While both mortality estimates and a plague dummy are used, there is doubt about the quality and comparability of mortality figures (Roosen & Curtis (2018), Dittmar & Meisenzahl (2019)), pointing to the plague dummy as being more robust. There is no evidence for demographic heterogeneity in mortality (Raster (2023), Alfani & Murphy (2017)) Morever, there is no evidence

[5] Appendix Table A.C.15 lists plagued region across all of Europe, even outside my sample.
[6] See Appendix Table A.B.3.

that the plague was spread by trading specific commodities.[7]

Figure 1.2.1 shows the 55 out of 594 regions that were plagued. As only seven plagued regions were additionally directly besieged, I expect the destruction of physical capital to play a minor role. Appendix Table A.B.5 shows that plagued cities were not significantly larger before the plague. I also show that plagued cities were not significantly smaller by 1750.[8] To test regional and not only urban population recovery, I show in Appendix Table A.B.6 that plagued Swedish parishes had recovered their populations by 1750. In Appendix A.B.3, I discuss the contributions of increased fertility, lowered mortality, and migration to this population recovery.

Figure 1.2.1: Plagued regions



Notes: **Green circles:** plague. **Beige circles:** no plague.

The main analysis identifies treatment as a plague outbreak in a region. As a robustness check, I incorporate plague outbreaks across all of Europe (Appendix Table A.C.15 and Appendix Figure A.C.21). I construct an indirect plague treatment variable that assigns higher weights to closer by and larger plagued regions. All details can be found in Appendix A.C.5. In Appendix Table A.C.16, I show that the indirect plague treatment variable is never significant, and that the main results established for the direct plague treatment go through as before. Therefore, I

[7]Frandsen (2009) mentions that in some places, textiles of those who perished were burnt as they were believed to be contagious. However, the finding in Appendix Figure A.E.13 that relative demand did not shift into textiles after the plague speaks against the salience of this anecdote.

[8]Given low urbanisation rates, I consider urban population recovery indicative of regional population recovery.

continue to focus on the direct effects of the plague.

1.2.2 The Little Ice Age

The period after the plague coincides with warming at the end of the Little Ice Age (Waldinger (2022)). I control for temperature change when recovering agricultural productivity: warming in a cold region with long winters raises agricultural productivity and lengthens growing seasons.[9] I use temperature data (Xoplaki (2005), Luterbacher et al. (2004)) and match regions to the four closest observations, weighting by the inverse distance to these grid points. I define growing seasons as spring and summer.

1.2.3 Serfdom

Appendix Figure A.E.11 shows regions by their serfdom status. While Denmark is the only country in this area that re-introduced serfdom in 1733 (Gary et al. (2022)), Raster (2023) argues that labour coercion increased in Northern Estonia as a result of post-plague labour scarcity. My paper does not seek to disentangle the margins of labour coercion. Instead, I conceptualise serfdom as a mobility restriction and collect data on the presence of serfdom. While serfs accounted for less than 10% of the population in Estonia and 10-30% in Lithuania (Baten et al. (2016)), labour coercion will have restricted the mobility of a significantly higher share of the rural population. My main sources are Raster (2023) and Peters (2022) (see Appendix A.E.3).

1.2.4 The Soundtoll Data

A novel granular trade panel is constructed based on the Soundtoll data (see also Marczinek et al. (2025)). I report key descriptives in Appendix Table A.A.1. Virtually all trade between the Baltic and North Seas is recorded.

Classification of Goods

I map 143,855 cargo descriptions in Old Danish to 227 goods and to five sectors: labour-intensive agriculture, capital-intensive agriculture, labour-intensive manufacturing, capital-intensive manufacturing, and remaining unclassified goods. As a robustness exercise, I use Claude's Opus 4.6 model to classify cargoes. This simple exercise produces the same goods classification as my

[9]The documented shift into capital-intensive agriculture actually runs counter to climate warming. Under warmer temperatures, regions were likely to reverse their shift into pastoral farming that had occurred during the Little Ice Age (Degroot et al. (2022), Degroot (2018)). This reversal should go in the opposite direction of what I document below.

manual efforts for 84% of traded value, and for 96% of traded value after some inconsequential differences[10] are accounted for (Appendix A.A.1).

In Section 1.4, I argue that the differentiation between labour- and capital-intensive goods is crucial, and that land intensities are not a driving factor.[11] I classify goods by recognising that capital-intensive goods require a high baseline amount of capital, contained in buildings, furnaces, machinery, tools, carts, and animals. Labour-intensive goods, on the other hand, are dominated by labour-intensive tasks (Heblich et al. (2025)) and limited in the degree to which labour can be saved.

Labour-intensive agriculture is arable farming. Such farming certainly uses some capital in the form of implements such as carts and ploughs. However, arable farming is dominated by the numerous labour-intensive tasks it entails, such as ploughing, sowing, weeding, harvesting and threshing (Heblich et al. (2025)). Pastoral farming, mining, and processed foods are capital-intensive agriculture. The infrastructure and tools required to mine and process foods make them capital-intensive. For pastoral farming, the importance of capital is well-established, with livestock itself classified as capital (Broadberry et al. (2014), Allen (2005)). Livestock require capital-investments for grazing (fencing, drainage, and roads) and for milking and slaughtering (buildings and tools) and otherwise graze freely in fields, requiring little labour input.

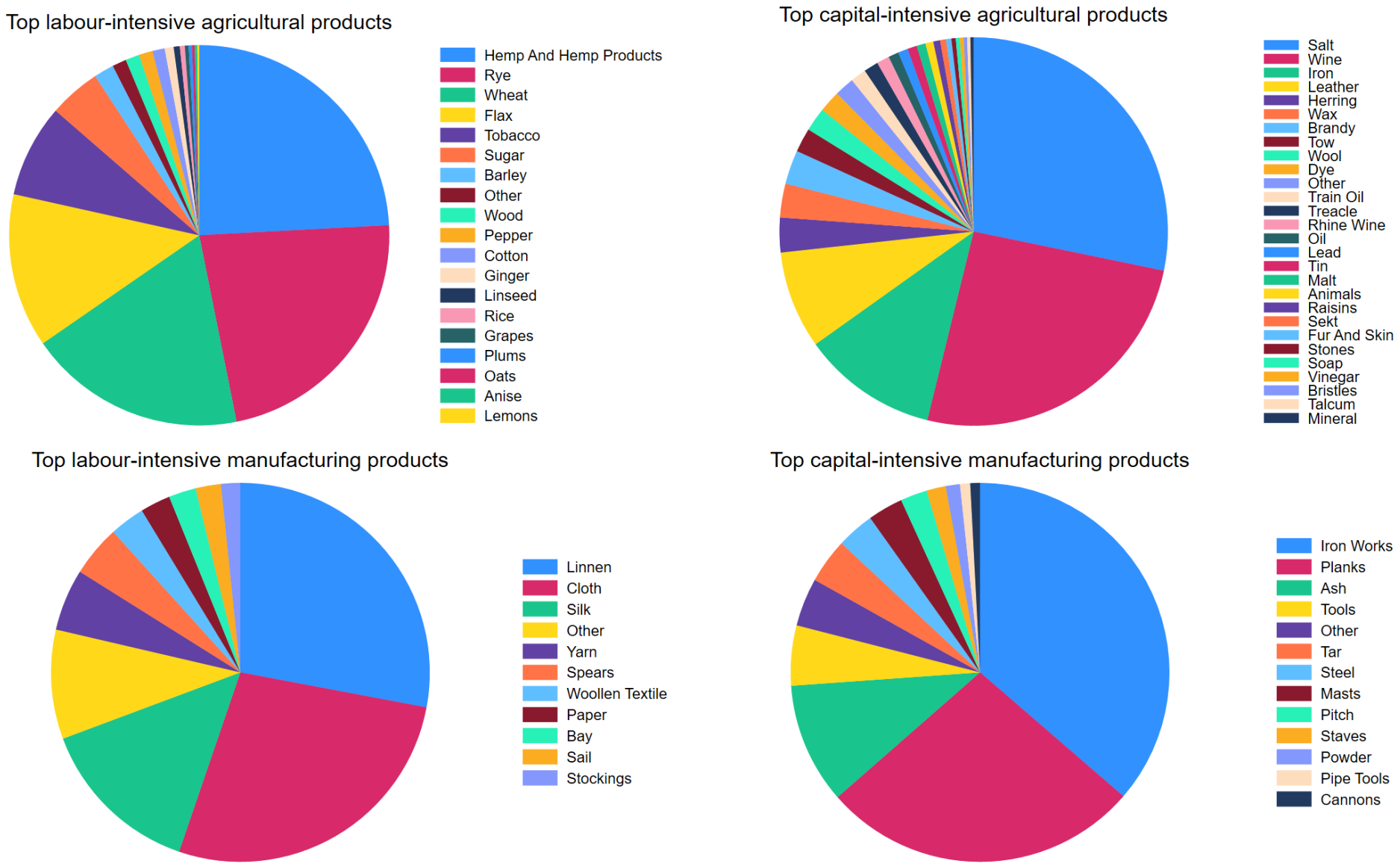
In manufacturing, I classify textiles as labour-intensive. Broadberry & Gupta (2009) and Pomeranz (2006) discuss how traditional production methods in textiles were very labour-intensive, with tasks such as spinning, weaving, and dyeing requiring high labour inputs. Textile production in the early 18th century Baltic region will have employed traditional production methods as these were agrarian societies before the Industrial Revolution. Iron and metal works and ship building, on the other hand, are naturally more capital-intensive in that they use furnaces, buildings, docks, and tools.

Figure 1.2.2 shows the most common goods for each sector by value between 1668 and 1750. Appendix Table A.A.2 ranks the top 30 goods, accounting for 90% of traded value during this period, and shows their assignments. I run a robustness exercise in which I reassign four debatable goods: wine and tow to labour-intensive agriculture and planks and ash to labour-intensive manufacturing. This does not change my results.

[10] These are iron works/iron, planks/wood, goods/mercery, cloth/sail, meat/beef.

[11] I find in Appendix Figure A.C.5 that there is no relative shift into land-intensive commodities (arable crops, livestock and animal products, plantation goods, raw forestry) after the plague.

Figure 1.2.2: Goods by sector



Notes: This figure shows the most common goods per sector by overall traded value. The 2% least common goods have been summarised as 'other' for legibility.

Classification of cities, duties, currencies, and units

The Soundtoll data record origins and destinations at the level of ports. I map 90,737 original port names to 3,085 unique port identifiers and assign them into one of 56 areas. Areas are mostly equivalent to NUTS 2 regions; for the United Kingdom, France, Germany, and Spain, they are NUTS 1 regions to ensure similarly sized areas and to avoid areas with a single port (Appendix A.B). However, my estimates are robust to grouping ports into NUTS 1 or NUTS 2 regions. Only ships passing the Sound are recorded. As such, I only observe trade between the Baltic Sea and the North Sea, but none within either. I confirm my main empirical findings in trade data for Hamburg (Appendix Table A.C.14) for which I see trade with all partners. I also present simulation results in Appendix Figure A.D.1 on a simulated productivity increase. Both in the hypothetical full and in the actually observed sample the productivity increase is recovered. In both cases the estimates are centred around the simulated effect, suggesting no bias.

Details on tolls and duties are given in Appendix A.A.3. I follow the literature in interpreting the proportional duty as a proxy for value (Waldinger (2022). Additionally, I verify that duty rate variation does not affect my findings (Appendix A.A.4). Currency concerns, as elaborated in Appendix A.A.2, are minor as 99.76% of trade transactions are carried out in Danish coins. I also use shipments' weights to construct unit prices (Appendix Table A.A.5). While there are reasons to doubt the accuracy of reported tolls and to suspect corruption, a truth-telling mechanism was in place to mitigate incorrect reporting: the Crown could purchase goods at the stated value to discourage underreporting.

1.3 Trade

This Section presents novel descriptives on trade following a plague outbreak. Labour scarcity[12] increases wages (Jedwab et al. (2022)), making capital relatively cheaper.[13] One needs to classify goods as labour- or capital-intensive. For agriculture, I classify arable farming as labour- and pastoral farming as capital-intensive (Heblich et al. (2025), Broadberry et al. (2014), Allen (2005)). For manufacturing, I distinguish labour-intensive production of textiles from capital-intensive metal works, ship building, and machinery-intensive production (Broadberry & Gupta (2009)).[14] Details on goods classifications can be found in Section 1.2.4.

Three adjustments could appear. First, an adjustment *within* sectors. As long as factor

[12]The terms scarcity and shortage are used interchangeably (Caldara et al. (2024)). I speak of scarcity throughout, as labour became scarcer compared to capital.

[13]In Appendix Table A.E.1, I follow Keller et al. (2021) in inferring interest rate from monthly grain price variation, and find that interest rates do not rise after the plague.

[14]I assume unclassified goods as labour-intensive; my results are unchanged when dropping them.

prices are not equalised across regions, a higher capital-to-labour ratio will be employed in the production of goods. In the case of arable farming, this might take the form of metal ploughs or drainage. In manufacturing, machinery and tools reflect a common way of increasing the capital-to-labour ratio. This adjustment *within* sectors is difficult to observe in trade data, so I turn to production data below.

The second adjustment is a compositional effect *across* sectors and as such observable in trade data. Labour scarcity should induce regions to increase their production of capital-intensive goods relative to labour-intensive goods. This adjustment *across sectors* is predicted by the Rybczynski theorem as part of the Heckscher-Ohlin theory of trade. At constant relative goods prices, it states that the rise in the endowment of capital, compared to labour, leads to a more than proportional expansion of output in capital-intensive sectors. This change in regional relative factor endowments does not lead to changes in regional factor prices as long as the region is sufficiently small. This prediction of factor price equalisation is rejected by the data, which is one of the reasons why I propose a Ricardian model.[15] The Ricardian model also predicts an adjustment *across sectors*, as labour is assumed immobile across regions, and can further rationalise an overall export expansion through productivity growth.

Third, labour scarcity may lead to productivity growth. If it does not, export would decline overall as wages rise; the first two adjustments would lead to changes in the capital intensity and composition of exports. If labour scarcity does lead to productivity growth, one should observe both a compositional change and an expansion of exports.

I expect the compositional adjustment across sectors to become visible in trade data almost immediately as producers respond to higher wages as soon as mortality reduced labour supply. It may be useful to consider the scope of this shock as the median mortality rate was 36%. Productivity growth, however, should take several years to materialise as producers experiment with new production methods and best practices take time to spread throughout the economy (Juhász, Squicciarini & Voigtländer (2024)). Furthermore, trade volumes during the second half of the war were severely depressed: after 1709, the Danish navy blocked ships; the Russian navy raided many Baltic coastal areas and intercepted vessels. This reduced the number of passages at the Sound (Appendix Figure A.C.1). I collect detailed data on Russian naval raids and blocked ports and exploit the fact that the Danish navy was operating its blockade out of the island of Christiansø. In Appendix Table A.C.2, I show that halving a port's distance to Christiansø is associated with 27-34% less ships sailing from that port during the blockade, controlling for

[15] Trefler (1993) reconciles the same empirical rejection of FPE with factor content theory by introducing factor-augmenting productivity differences. I will return to this point in Section 1.6.

other direct blockades, Russian raids, and a host of fixed effects. Crucially, this adverse effect on trade does not vary by a port's plague status. This increased risk of shipping thus affected both plagued and unplagued ports and limits our ability to distinguish their export composition and levels after the plague. Therefore, trade data from the first decade after the plague are relatively uninformative, and most effects become visible only after the peace treaty of 1721 restored shipping routes. This is confirmed by studying the export volumes and composition of the main plagued port on the North Sea, where shipping was not disrupted: Hamburg's trade (see Appendix Table A.C.14) shows an effect on impact.

Identification

I provide evidence to rule out that the plague was endogenous to trade. In the literature on the Black Death, trade has been proposed to have propagated the plague; for this outbreak, however, armies have been argued to have spread the plague and rarely directly targeted cities. The main reason for this is that during the period when trade could have spread the plague, the Danish blockade severely reduced trade volumes, which supports my claim that this particular plague was exogenous to trade. There is no evidence of the plague as a biological weapon. In a balance test, to-be-plagued cities do not differ in population size before the plague (Appendix Table A.B.5). For trade, Table 1.3.1 shows that a future plague dummy does not predict a different volume, composition, or number of exports. However, to-be-plagued regions are likelier to export at all. Yet, Table 1.3.1 shows that to-be-plagued ports do not differ in their trade growth. The level difference will be absorbed below with origin fixed effects and the analysis will focus on trade growth. Balancedness in this regard is confirmed in that I tend to not find any significant pre-trends in event studies.

Table 1.3.1: Pre-plague balance checks

	Log Exports	Export Growth	Log # of Exported Goods	Export Probability
	(1)	(2)	(3)	(4)
To be Plagued	0.588	-0.055	0.221*	0.090***
	(0.427)	(0.064)	(0.120)	(0.018)
To be Plagued x Capital-Int.	-0.437	0.080	0.043	-0.048***
	(0.367)	(0.063)	(0.107)	(0.006)
Fixed Effects:				
– Destination x Year	✓	✓	✓	✓
Years	1668-1708	1689-1708	1668-1708	1668-1708
Observations	15,826	47,160	31,604	496,920

Notes: Standard errors clustered at the origin level in parentheses. * p <0.10, ** p <0.05, *** p <0.01. The dependent variables are log export volumes (column 1), annual export volumes divided by total exports between 1668 and 1688 (column 2), the log number of exported goods (column 3), and a dummy for any exports (column 4). The independent variable is a plague dummy, equal to one for to-be-plagued regions, and this dummy's interaction with a dummy for a capital-intensive sector. Years are restricted to before 1708 to exclude any plague years. Additional controls are annual growing season temperature, latitude, and longitude.

Fact #1: Capital-intensive exports increase more than labour-intensive exports.

The plague induced labour scarcity, which will have translated into higher wages and an increased capital intensity of production. The plague's effects should thus differ by sectors' factor intensities. To test this, I estimate an event study using OLS on log exports, weighted by exports in levels.[16] In particular, I estimate:

$$T_{ikt} = \sum_{l=-20\setminus -1}^{30} \beta_l^f \cdot 1(K_{it} = l) + \alpha_{ik} + \alpha_{kt} + \gamma_f x_{it} + \epsilon_{ikt}, \tag{1.3.1}$$

where i are origin ports, k one of 5 sectors, f is factor intensity (capital- or labour-intensive), and t are years. T_{ikt} are (log) exports.[17] K_{it} is the time difference to the plague, α_{ik} are origin-sector fixed effects, α_{kt} sector-time fixed effects, and x_{it} are annual growing season temperatures. I am particularly interested in the difference between β_l^C and β_l^L. Throughout, standard errors are clustered at the origin-level.

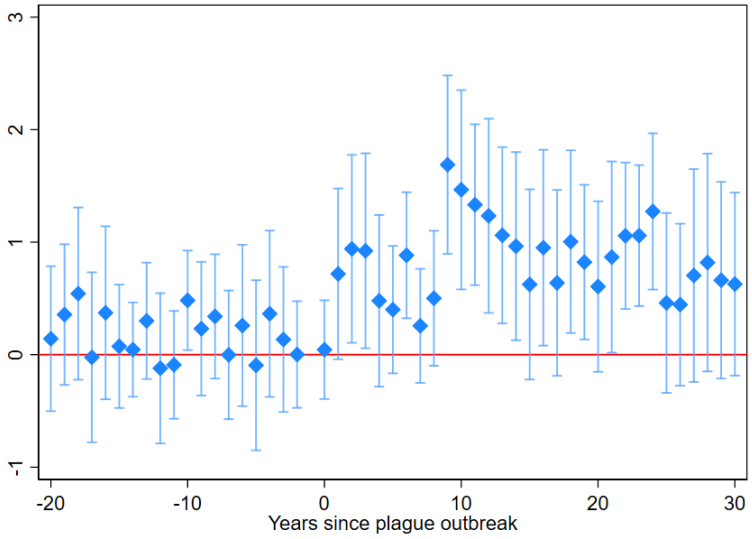
Figure 1.3.1 shows the difference between β_l^C and β_l^L: after the plague, capital-intensive exports expand relative to labour-intensive exports. The estimates imply a 10pp increase in the share of capital-intensive exports, compared to a pre-plague mean of 6.6%.[18]

[16] This recovers OLS estimates similar to PPML estimates. Mayer et al. (2019) and Head & Mayer (2013) show that PPML places more weight on high trade flows; weighting OLS on log trade by levels thus brings OLS close to PPML results (Mayer et al. (2019)).

[17] Throughout, these are the value of exports. 99.76% of transactions were carried out in Danish coins.

[18] As discussed above, the shock becomes particularly visible after a decade, when war-induced shipping disruptions unrelated to the plague had been resolved (Appendix Figure A.C.1 and Appendix Table

Figure 1.3.1: Capital-intensive exports expand relative to labour-intensive exports



Notes: Estimation of equation 1.3.1. Standard errors clustered at the origin level.

In the Appendix, I present further results. Appendix Figure A.C.2 shows that an alternative assignment of four debatable goods produces almost identical results. Appendix Table A.C.3 shows PPML and OLS results in a DiDiD set up. Appendix Table A.C.4 shows that army proximity and sieges are not associated with a shift into capital-intensive exports. This also serves as a placebo test: for unplagued regions, army proximity and sieges do not predict changes in export composition. Appendix Figure A.C.3 confirms the above finding using the imputation estimator in Borusyak, Jaravel & Spiess (2021). Appendix Table A.C.5 shows that the relative expansion of capital-intensive exports occurred both in manufacturing and in agriculture. Further, Appendix Figure A.C.5 shows that no shift into land-intensive exports takes place after the plague. I analyse the plague's effects on the composition of exports at the goods level in Appendix Figure A.C.4.

Directly on production, I show that after the plague Swedish farms switched out of labour-intensive rye and barley production and into raising calves and foals. These effects are stronger the closer a farm is located to a plagued city (Section

This shift into capital-intensive exports outlasts the shock to factor proportions by at least half a century. In Appendix Tables A.B.5 and A.B.6, I show that by 1750 plagued cities and parishes were no smaller than unplagued ones. Given population recovery, factor proportion theories of trade would suggest that the shift across sectors should be reversed. To the contrary, Appendix Figure A.C.6 shows that the shift into capital-intensive exports persists for almost 90

A.C.2).

years after the plague. Capital intensity thus remains significantly elevated despite population recovery.

Fact #2: Plagued regions capture larger shares of destination markets.

I estimate a second event study incorporating the gravity structure of trade:

$$T_{ijt} = \sum_{l=-20\setminus-1}^{30} \beta_l \cdot 1(K_{it} = l) + \alpha_{ij} + \alpha_{jt} + \alpha_{at} + \gamma x_{it} + \epsilon_{ijt}, \tag{1.3.2}$$

where T_{ijt} is one of several outcomes for trade from origin i to destination j in year t. K_{it} is the time difference between the current year and the plague, α_{ij} is a bilateral fixed effect, and α_{jt} are destination-time fixed effects absorbing the demand side. ϵ_{ijt} is the error term. α_{at} are area-time fixed effects.[19] These permit for areas' different time trends and thus permit heterogeneity by area. In particular, they absorb area-wide changes in market access due to territorial changes, such as Estonia and Latvia becoming a part of Russia, and area-wide shocks such as famines. x_{it} are time-varying origin controls, in particular growing season temperature, longitude x year, and latitude x year. β_l are the outcomes of interest.

For Fact #2, T_{ijt} is either the volume of exports from i to j in t, or the export share that origin i captures in destination j: $T_{ijt} = \frac{x_{ijt}}{\Sigma_{i \in I} x_{ijt}}$.[20] This formulation of T_{ijt} as the share of exports from origin i in destination j in year t motivates the Ricardian model below, which predicts trade shares. For results on the extensive margin and on new exports, T_{ijt} is the number of goods exported from i to j in t.

Contrary to concerns of an export contraction, Figure 1.3.2 shows that plagued cities increase their export shares by about 1pp, which is close to the pre-plague mean export share of 0.9%. Appendix Figure A.C.7a also shows an expansion of export volumes.

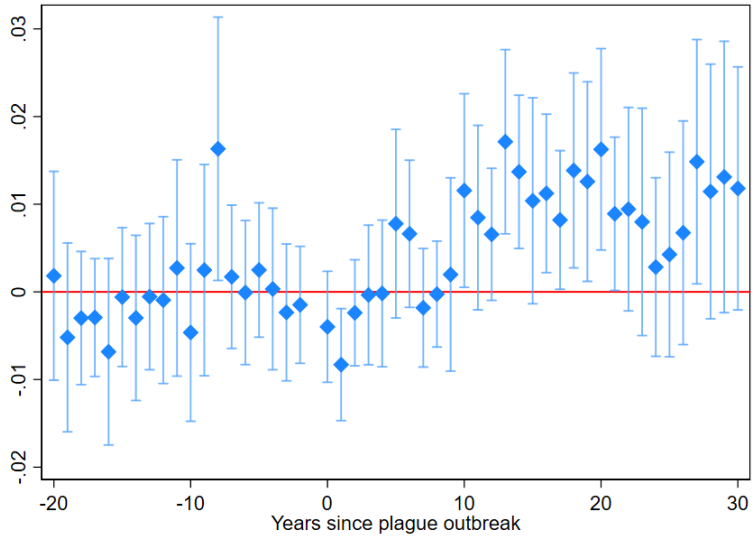
The export expansion begins shortly after the plague, with the first significant point estimates for export volumes three years later. Ten years after the plague, these regions capture significantly larger market shares in destination markets. This time difference reflects the above-discussed war-related disruptions to shipping in the Baltic Sea, which were unrelated to the plague (Appendix Figure A.C.1 and Appendix Table A.C.2).

Notably, these results reject the hypothesis that trade simply bounced back. There is a

[19] Areas are displayed in Appendix Figure A.B.1. For most of Europe, these are equivalent to NUTS2; for the UK, Germany, France, and Spain, they are NUTS1. This is to ensure similarly sized areas and to avoid areas with a single port (Appendix A.B).

[20] I set market shares of i in j to 0 whenever j shows no imports in my data. As the data capture only trade between the North and Baltic Seas, I thus assume that every region has some imports, even if unobserved to me.

Figure 1.3.2: Market shares expand



Notes: Estimation of equation 1.3.2 on i's share of j's imports, $T_{ijt} = \frac{x_{ijt}}{\sum_{i \in I} x_{ijt}}$. Standard errors clustered at the origin level.

decade of peace during which to-be-plagued ports do not display different export patterns. Appendix Table A.C.6 presents results from a difference-in-differences set up using OLS and PPML, excluding the war years of 1700 to 1721. Appendix Figure A.C.8a extends the pre-period to 30 years and again finds no pre-trends. To more flexibly control for temperature changes, Appendix Figure A.C.11a additionally controls for temperatures with lags of 10 and 20 years. Appendix Figure A.C.8b constructs market shares from cleaned underlying value correcting for duty rates (Section 1.2.4 and Appendix A.A.4). Appendix Figure A.C.9 finds very similar results with NUTS 1-time and NUTS 2-time fixed effects instead of the proposed area-time fixed effects.

Appendix Figure A.C.11b implements the estimator by Borusyak, Jaravel & Spiess (2021). As the summation of ATEs may involve negative weights, I check and find that all weights are positive and conclude that the standard estimator is appropriate (de Chaisemartin & D'Haultfoeuille (2022)). Similarly, I address staggered treatment timing by implementing the decomposition in Goodman-Bacon (2021). I find that more than 99.1% of the identifying variation comes from comparisons between never-treated and eventually-treated ports, with less than 0.9% coming from comparisons across treatment timings. My estimates are therefore virtually unaffected by potential bias of this kind.

Appendix Figure A.C.10 decomposes the intensive margin results into sectors. Two decades after the plague, I find that capital-intensive manufacturing and agriculture contribute each a quarter to the overall export expansion, with the remaining half accruing to labour-intensive agriculture. Labour-intensive manufacturing plays virtually no role. The outsized role of labour-

intensive agriculture is not surprising as it makes up 55% of traded value between 1668 and 1750. Aggregating sectors by factor intensity instead, Appendix Figure A.C.12 shows that both saw export expansions. These results choose T_{ijft} to be the volume of exports by factor intensity. At a more granular level, Appendix Figure A.C.13 presents results at the goods level.

The increase in destination market shares may be indicative of productivity growth as plagued regions became more competitive. This export expansion outlives population recovery which had been achieved by 1750. Appendix Figures A.C.14a and A.C.14b show that up to 80 years after the plague, destination market shares and export volumes are still larger in previously plagued regions.

Extensive Margin: Plagued regions export a larger variety of goods.

The extensive margin plays an important role in the export expansion: Appendix Figure A.C.7b shows that plagued regions increased the number of goods they export. Through the lens of a Ricardian model (Eaton & Kortum (2002)), the extensive margin expansion drives the results on market shares. Plagued regions becoming the most competitive suppliers of more goods to destination markets is therefore an indication of productivity growth and observationally consistent with an export expansion along the intensive margin. Appendix Figure A.C.15 decomposes the extensive margin expansion by sector. Appendix Figure A.C.16 shows these results overall for capital- and labour-intensive sectors. Appendix Figure A.C.17 shows results at the level of goods.

1.3.1 Other Robustness Results

I provide results from a PPML panel regression on trade levels in Appendix Table A.C.9. The results show that trade increased after one side suffered a plague outbreak. Considering the set of fixed effects, however, this result draws on different variation from my main result. Further, Appendix Table A.C.10 shows results for the probability of a region exporting in a given year. After a plague outbreak, regions' probability to be active exporters increases significantly. I present heterogeneity results in Appendix Table A.C.11. While the effects are larger for regions that exported more before the plague, they are significant even for regions that had no exports before the plague. Appendix Table A.C.13 shows that the presence of a besieged city does not affect the post-plague trade findings. Finally, I check if the specification of treatment is sufficient in focusing on plague outbreaks in regions themselves. In Appendix A.C.5, I describe how I construct a continuous plague treatment variable that captures plague outbreaks across

all of Europe. I show in Appendix Table A.C.16 that the indirect treatment variable is never significant.[21] Further, my findings on the direct plague treatment are robust to including the indirect plague treatment.

1.3.2 Mechanical Explanations

I argue that my findings are not driven by mechanical forces. First, the observed export expansion might reflect a reallocation of exports from the Baltic Sea to the North Sea. Heterogeneity across sectors speaks against a simplistic reallocation. Further, this raises the question of how plagued exporters became competitive in new markets where they competed against existing exporters. The combination of new markets and longer distances implies that regions in the Baltics would have had to overcome the cost of distance and higher wages through higher productivity to be competitive.

To more rigorously address this concern, I show simulation results in Appendix Figure A.D.1 for a simulated productivity increase and its effect on trade flows. Both in the hypothetical full and in the actually observed sample, where only trade passing the toll station is registered, the productivity increase is reflected in significantly higher exports. For both samples, the point estimates are centred around the simulated effect, suggesting no bias.

Territorial changes may also affect the results. Sweden lost control of modern-day Estonia, Latvia, and parts of Northern Germany. This affected these regions' market access to Sweden, but also Swedish ports' market access to the lost territories. This concern is addressed by including area-time (e.g. Estonia-year) fixed effects. The variation I draw on is between ports within Estonia, all of which underwent changes in market access. These fixed effects thus account for territorial changes. Relatedly, the late 17th and early 18th century saw several particularly harsh winters and led to famines. As little localised information is available on famines, I use area-time fixed effects to absorb area-wide famines. In Appendix Figure A.C.11a, I additionally control for growing season temperatures and for temperatures with a lag of 10 and 20 years to control for harsh winters.

A second mechanical explanation concerns the role of harbours as export hubs. If a harbour exports hinterland production, a share of these products will be consumed in the port itself. Following a plague outbreak, local consumption may drop, and, assuming constant hinterland production, exports should rise.[22] To test whether local demand contraction in export hubs

[21] This also holds for imports. Results available on request.

[22] I challenge the notion that the hinterland's production would not have been affected. The plague was spread by moving armies, which plundered the hinterland and disrupted production.

drives my results[23], I show in Appendix Figure A.C.20 that exports shipped to plagued regions increase. The mechanical explanation predicts a decline in imports, however. Further, Appendix Table A.C.11 shows that also the smallest harbours that are unlikely to have served as export hubs see an export expansion.

Relatedly, one could argue that my findings are driven by exporter-specific declines in trade costs related to the plague or to a higher capital intensity of exports. After the plague, harbour operations may have become more efficient, and capital-intensive goods may be faster to load onto ships. To address this issue, I count the number of days between two appearances at the Sound of the same captain (identified by full name and home town) on return trips, where the first leg's destination overlaps with the second leg's origin. This is an exporter-specific measure of both travel days and handling time at harbours, where captain-specific and origin-time determinants of days at sea are absorbed by fixed effects. In Appendix Table A.C.12, I show that travel and loading times are unaffected by either the plague or by capital intensity. Thus, trade costs are unrelated to the shock.

Finally, I provide evidence in Appendix Table A.C.14 against both concerns - selection and mechanical reallocations - by drawing on data by Baasch (1929) that cover all of Hamburg's trade by commodity between 1702 and 1713 (Pfister (n.d.)). Exports are recorded with all of Hamburg's trading partners. This provides more detail than the Soundtoll data, which are restricted to trade between the Baltic and North Seas. I find evidence that exports expand after the plague and shift into capital-intensive commodities. This dispels concerns that Baltic exporters simply entered into markets that plagued Hamburg had left. Similarly, this finding confirms the export expansion and shift into capital-intensive exports for one port for which I have export data for all destinations. As Hamburg was unaffected by naval warfare and sees effects on impact, this suggests that the time lag discussed above is indeed driven by war. In sum, mechanical forces are unlikely to drive my findings.

1.4 A Ricardian Model

I propose a Ricardian model to explain my empirical findings. In Malthusian models with fixed productivity, an adverse labour supply shock would be followed by an export decline as wages increase. If only Heckscher-Ohlin forces were at play, one would observe a reallocation to capital-intensive production but no overall increase in exports. I argue more extensively against pure

[23]In national accounting terms, exports rise if domestic absorption falls by more than income, such that production decreases less than consumption.

Heckscher-Ohlin forces in Section 1.6. Finally, factor theories predict factor price equalisation.[24] As all three are rejected by my findings, I build a Ricardian model with intra-industry variation in productivity (Donaldson & Hornbeck (2016), Costinot et al. (2011), Eaton & Kortum (2002)).

There are many regions, i=1,...,I, and two factors of production: labour L_{ik} and capital K_{ik}. There are five sectors indexed by k: labour-intensive agriculture (LA), capital-intensive agriculture (CA), labour-intensive manufacturing (LM), capital-intensive manufacturing (CM), and unclassified goods (U). Throughout, I am dropping the t subscript indicating time for simplicity.

Unique Spatial Equilibrium: Population recovery is an indicator of a unique spatial equilibrium (Davis & Weinstein (2002)). I use city population data by Buringh (2021) to show that by 1750, plagued cities and parishes had returned to their pre-plague growth paths (Appendix Table A.B.5 and A.B.6). While for a sufficiently small shock population may also recover in a NEG setting such as Krugman (1991), I argue that the median mortality rate of 36% is anything but small. Thus, the model has a unique spatial equilibrium.

Factors and Sectors: The model features labour and capital. Each sector has a labour share γ_k. Capital is assumed to be freely mobile across regions and sectors. For each sector k the production function takes the Cobb-Douglas form:

$$P(L_{ik}, K_{ik}) = A_{ik} L_{ik}^{\gamma_k} K_{ik}^{1-\gamma_k}. \tag{1.4.1}$$

One might wonder whether land ought to be included. I argue that land was abundant in the Baltics: Raster (2023) shows that population densities were so low that after the plague entire manors were abandoned; Allen (2003) proposes that serfdom in the Russian Empire was an institution to preserve the power of the owners of abundant land over scarce peasants. Therefore, I argue that labour was the constraining factor. Had land been a constraining factor, the increased relative supply of land should benefit the production of land-intensive goods; Appendix Figure A.E.5 finds no evidence in production data and Appendix Figure A.C.5 find no shift into land-intensive exports. Further, my findings are also present within manufacturing, where land intensities should play n role (Appendix Table A.C.5 and Appendix Figure A.C.12). Therefore,

[24] With more than two countries, Helpman & Krugman (1985) show that FPE continues to hold if all countries fall within the FPE set, which comes down to endowments not being too dissimilar. Schott (2003) shows that FPE is achieved for many countries if they are in the same diversification cone: as countries become more similar in endowments, FPE holds approximately also for n>2, and precisely when endowments are identical.

Table 1.4.1: Spatial Separation & Capital Intensity

		Labour-intensive		Capital-intensive
Region	City	Labour-int. manufacturing (LM)	$\gamma_{LM} > \gamma_{CM}$	Capital-int. manufacturing (CM)
	Hinterland	Labour-int. agriculture (LA)	$\gamma_{LA} > \gamma_{CA}$	Capital-int. agriculture (CA)

Notes: Assumptions on spatial separation, sectors, and factor intensities. γ_k denotes the labour share in each sector k.

I argue that the crucial comparison is between labour- and capital-intensive goods and do not model land as a factor.

I assume cities to produce manufacturing and hinterlands to produce agricultural goods. This assumption finds historical support in policies that restricted manufacturing to cities. In Sweden, the textile industry was restricted to cities, with rural producers supplying only the input goods; similarly, urban blast furnaces were to be supplied by rural pig iron (Magnusson (2007)). To ensure this separation, a 'town economic policy' was introduced in the 17th century. In a similar vein, Klein & Ogilvie (2015) describe how urban institutions hindered rural industrial production.

I define a region as a city and its hinterland. Larger cities on navigable rivers may boast larger hinterlands; all I assume is that this does not change over the study period. Swedish cities, for example, enjoyed trade monopolies over their surrounding area, which only changed in 1765 when the Commodity Act was withdrawn (Magnusson (2007)). All 1,425 cities in the data are coastal, and the average diameter of the hinterland is 25 kilometres. Table 1.4.1 summarises these assumptions.

Labour Mobility: With labour mobility across regions, wage differentials are eliminated through migration. One would have to observe an immediate population recovery under this assumption; however, I find that population recovery was achieved only by 1750. The literature on plague outbreaks finds significant wage increases (Jedwab et al. (2022)), implying local labour scarcity and the absence of factor price equalisation. The latter is supported by Davis & Weinstein (2001). Accordingly, I model labour as immobile across regions but mobile across sectors within a region. Plagued cities, where mortality rates were higher than in the hinterland, can recover via rural-to-urban migration. This is true in areas without serfdom. Under serfdom, however, peasants cannot move to cities and a wedge between urban and rural wages persists.

Workers are mobile within agriculture and within manufacturing in a given region, implying wage equalisation.[25] Let $w_i^{CM} = w_i^{LM} \equiv w_i^M$ and $w_i^{CA} = w_i^{LA} \equiv w_i^A$. Without serfdom, all four

[25]While serfs are not paid wages and cannot move between different agricultural employments, the assumption here is that their lords efficiently allocate them to equalise the marginal products of labour

sectors pay equal wages in a region. With serfdom, there are equalised wages within agriculture and within manufacturing. The wedge between urban and rural wages, ϕ_i, will be region-specific and depends on mortality differences between the two parts of a region. ϕ_i also captures language and skill frictions. Equation 1.4.2 summarises the assumptions on labour mobility. Appendix Figure A.E.12 shows that a wedge ϕ_i appeared in Denmark after the re-introduction of serfdom.

$$w_{iM} = \begin{cases} w_{iA}, & \text{without serfdom,} \\ (1+\phi_i)w_{iA}, & \text{with serfdom.} \end{cases} \tag{1.4.2}$$

Productivity: Each sector's output features in an infinite number of varieties, $\omega \epsilon \Omega$. Productivity A_{ik} is a random variable, drawn independently for each (i, k, ω, t) from a Fréchet distribution as in Donaldson & Hornbeck (2016):

$$F_{ik}(z) = 1 - exp(-A_{ik} z^{-\theta}), \tag{1.4.3}$$

where $A_{ik} > 0$ and $\theta > 1$. A_{ik} captures fundamental productivity of region i in sector k, and encompasses productivity-affecting fundamentals that pertain to all producers in that region and sector. θ reflects intra-industry heterogeneity.[26] In this model, productivity is therefore total factor productivity and not factor biased (Acemoglu (2002)).[27]

Capital Intensity: Producers take the variety-level draw $A_{ik}(\omega)$ as given when allocating labour and capital. Factor prices equal marginal products, so the production of each variety ω becomes more intensive in capital after a wage increase:

$$\frac{K_{ik}(\omega)}{L_{ik}(\omega)} = \frac{1-\gamma_k}{\gamma_k r} w_{ik}.$$

The capital-to-labour ratios rise also overall, with $K_{ik} = \int_{\omega \in \Omega} K_{ik}(\omega)$. The model therefore predicts the shift *within* sectors into capital-intensive production. The strength of this shift depends on the elasticity of substitution between factors: this equation is an artefact of the Cobb-Douglas production function, where the elasticity of substitution between capital and labour is 1. Alternatively, production functions could be linear (with an infinite elasticity of substitution) or Leontieff (with an elasticity of substitution of 0). However, a linear production

between labour-intensive and capital-intensive agriculture.

[26] Costinot et al. (2011) elaborate on the implications of assuming $\theta_k = \theta \forall k$.

[27] Davis & Weinstein (2001) argue that a model of Hicks-neutral technical differences is best suited to explain global production structures.

function has empirically unsupported corner predictions for factor demand; further, the rising post-plague capital-to-labour ratio (see Section 1.5) speaks against a Leontieff function.

Marginal Costs: The marginal cost of production is given by:

$$MC_{ik}(\omega) = \frac{(w_{ik})^{\gamma_k} r^{1-\gamma_k}}{A_{ik}(\omega)}. \tag{1.4.4}$$

This is a model of constant marginal costs. My finding relates to the theory of venting out (Almunia et al. (2021)), in which a domestic demand slump leading to an export boom is rationalised by non-constant marginal costs of production. In Section 1.6, I argue that one can better rationalise my findings without assuming non-constant marginal costs.

Trade Costs: I assume the standard iceberg form. Trade costs are symmetric, $d_{ijk} = d_{ji,k}$. The no-arbitrage condition $d_{il,k} \leq d_{ijk} d_{jl,k}$ is also imposed. Trade costs may vary by sector and time, thus subsuming any potential tariffs.

Market Structure & Prices: Markets are perfectly competitive and consumers in j purchase a variety ω from its cheapest supplier location i:

$p_{jk}(\omega) = \min_{1 \leq i \leq I} c_{ijk}(\omega)$, where $c_{ijk}(\omega)$ is assumed to be strictly positive. Following Redding & Venables (2004) and Donaldson & Hornbeck (2016), consumer market access and firm market access are defined as the price indices:[28]

$$CMA_{jk} = (P_{jk})^{-\theta} = \chi_k \Sigma_{i=1}^{I} A_{ik} (w_{ik})^{-\gamma_k \theta} d_{ijk}^{-\theta}, \text{[29]} \tag{1.4.5}$$

$$FMA_{ik} = \chi_k \alpha_k \Sigma_{j=1}^{J} d_{ijk}^{-\theta} (CMA_{jk})^{-1} Y_j. \tag{1.4.6}$$

Wage Rate & Labour Force: Wages are determined sectorally. Combining market clearing conditions and plugging in firm market access yields:

$$w_{ik} = \zeta_k \left(\frac{A_{ik} FMA_{ik}}{L_{ik}} \right)^{\frac{1}{1+\gamma_k \theta}}. \tag{1.4.7}$$

[28] Note that the existence of these CES price indices requires the assumption $\sigma_k < 1 + \theta$.

[29] $\chi_k = \left(\Gamma(\frac{\theta+1-\sigma_k}{\theta}) \right)^{\frac{-\theta}{1-\sigma_k}} r^{-(1-\gamma_k)\theta}$; $\zeta_k = \left(r^{1-\gamma_k} \right)^{\frac{-\theta}{1+\gamma_k \theta}}$

Preferences: The representative consumer has a two-level utility function, where the upper tier is Cobb-Douglas and the lower tier is CES. α_k are the Cobb-Douglas weights in the upper tier and $\sigma_k > 1$ is the elasticity of substitution between differentiated varieties in the lower tier. I assume $\sigma_k < 1 + \theta$. A continuum of these differentiated varieties ω are consumed. I will later follow Fieler (2011) in introducing non-homotheticity.

Trade Flows: Trade flows take the gravity form:

$$X_{ijk} = \frac{\chi_k A_{ik} (w_{ik})^{-\gamma_k \theta} d_{ijk}^{-\theta}}{CMA_{jk}} \alpha_k Y_j. \tag{1.4.8}$$

Region i exports more to region j in sector k if it has a higher productivity, A_{ik}, lower trade costs, d_{ijk}, or lower wages, w_i. Region i further exports larger volumes to j if region j has low consumer market access in a sector, which means low competition for i.

Equation 1.4.8 can be used to recover productivity growth. Adding a time dimension to the procedure in Costinot et al. (2011), I estimate using PPML:

$$X_{ijkt} = exp(\alpha + \delta_{ijt} + \delta_{jkt} + \delta_{ikt}) \times \epsilon_{ijkt}, \tag{1.4.9}$$

where all δ parameters are fixed effects and ϵ_{ijkt} is the error term. The denominator in equation 1.4.8 is absorbed by δ_{jkt} fixed effects, which also absorb inward multilateral resistance (Fally (2015), Anderson & van Wincoop (2003)). There is no structural outward multilateral resistance term. However, wages absorb the general equilibrium feedback that would show up in a structural outward multilateral resistance term. The critical assumption is that wages are common across sectors and differenced out (see Appendix Section A.D.1). In this case, there is no outward multilateral resistance contamination in the structural sense, and δ_{ikt} recovers relative productivity growth.

1.5 Mechanism

I propose capital deepening-driven productivity growth to explain plagued regions' export success. In response to labour scarcity (Step #1), producers increase the amount of capital they use relative to labour (Step #2). While not internalised by producers, a sector-wide increase in the capital-to-labour ratio increases productivity (Step #3).

Importantly, I do not endogenise productivity. The literature recovering productivities requires the assumption of variety-level productivity draws from a given distribution (Costinot et al.

(2011)) and views the distribution's scale parameters as fundamentals. Origin-sector variation in trade data has been projected on observables to assess the determinants of productivity (Chor (2010)). Following this literature, I relate recovered productivity growth to the plague as a shock to factor proportions while absorbing other determinants with fixed effects.

I propose a reduced-form relationship to lend economic meaning to sectoral productivities. While producers take the variety-level draw $A_{ik}(\omega)$ as given, I suggest that the Fréchet distribution's scale parameters A_{ik} are a function of the capital-to-labour ratio $\frac{K_{ik}}{L_{ik}}$, a vector permitting permanent effects of shocks, H_{ik}, and sectoral, regional, and region-sector determinants x_i, x_k, x_{ik}:

$$A_{ik} = g\Big((\frac{K_{ik}}{L_{ik}})^{\beta_k}, H_{ik}, x_i, x_k, x_{ik}\Big). \tag{1.5.1}$$

The function g(.) is increasing in the capital-to-labour ratio, whose impact on productivity is scaled by β_k. Sectors differ in the degree to which increased capital can move productivities and technological barriers present the main hurdle. Appendix Table A.E.2 lists examples of capital deepening available in this period. In the production of wheat, e.g., capital deepening in the early 18th century had limited effects on productivity, as no modern farm machinery was available beyond limited tools. The scope for capital to increase agricultural productivity was severely limited in this period, with most productivity-increasing devices centuries old (Gallardo & Sauer (2018) and Coleman (1956)). Atack et al. (2019) concludes that few technologies were available before the 19th century to increase productivities in labour-intensive manufacturing. In capital-intensive iron and steel making, on the other hand, capital deepening and investments into more efficient furnaces more readily translate into productivity improvements.

H_{ik} accounts for lasting effects of shocks. While my model is static, I allow for permanent productivity effects of shocks. This formulation generalises the arguments in Krugman (1987), Juhász (2018), and Lane (2025). As there are no granular data for this period to test the elements of x_i and x_k, I absorb this variation with fixed effects. Origin-destination-time and destination-sector-time variation are absorbed when decomposing gravity equation 1.4.9. Below, I include origin-sector fixed effects to absorb the time-invariant part of vector x_{ik}. I also absorb area-sector-time variation with fixed effects, leaving only origin-sector-time variation within areas.

I then relate the remaining variation to the plague as a proxy for the increased capital-to-labour ratio. This simplifies the reduced-form equation above:

$$A_{ik} = \left(\frac{K_{ik}}{L_{ik}}\right)^{\beta_k}. \tag{1.5.2}$$

Given the variation which I propose to be driven by capital deepening, one alternative are human and social capital. If these were sectorally varying - in levels, they are absorbed by fixed effects - and respond to labour scarcity, the numerator in equation 1.5.2 may not capture only physical capital K_{ik}, but rather the sum of physical, human, and social capital, which plausibly all shift in the same direction after the plague. Goods that are intensive in physical capital may also requrie higher social or human capital. I thus need to check whether, holding factor intensity constant, there is a relative export shift into goods that intensively use human and social capital. In Section 1.6, I find no evidence for human or social capital depeening. Even when controlling for how intensively human and social capital are used in the production of goods, I document a strong relative shift into capital-intensive exports after the plague. Thus, I propose that equation 1.5.2 captures the most relevant reduced-form relationship between factor proportions and productivity.

Micro-Foundation

I propose a micro-foundation that produces equation 1.5.2 in Appendix A.E. I introduce technology choice as an endogenous component to productivity (Bustos (2011)). Producers may use a labour-intensive technology that requires low fixed costs. Alternatively, they may incur a higher fixed cost to produce under a capital-intensive technology. Examples of capital-intensive technology in this period are provided in Appendix Table A.E.2. This provides an endogenous boost to the exogenous Fréchet productivity draw, which I expect to be larger in capital-intensive sectors. The trade-off defines a threshold productivity draw beyond which the capital-intensive technology is employed. As wages rise, this threshold productivity drops; productivity and the capital-to-labour ratio rise. This defines an equation mirroring equation 1.5.2. The parameter β_k now reflects sectoral differences in the productivity boost conferred on users of the capital-intensive technology.

Further, learning by doing (Krugman (1987)), learning by exporting (Loecker (2013)), and cost discovery (Hausmann & Rodrik (2003)) can micro-found my proposed mechanism. Faced with a labour supply shock, plagued regions have to experiment with new factor proportions and new products. In doing so, they may reap the benefits of learning by doing or exporting and may discover their true comparative advantage. These all involve a higher capital-to-labour ratio and predict productivity growth. While I am unable to differentiate between these

competing micro-foundations, all of them would correctly be picked up as productivity growth below. Further, they all imply that comparative advantage patterns are not fully determined by contemporaneous fundamentals. In both senses, these micro-foundations are consistent with my proposed mechanism.

Step #1: The Plague Induces Labour Scarcity

The plague reduced urban (regional) populations by 36% (20%)[30] on average, which increases wages. According to equation 1.4.7, supposing $\gamma_k = 0.5$ and $\theta = 5$, a 20% decrease in employment increases sectoral wages by 6.6%. The same is true for regional wages, as sectoral employment adjusts so as to arbitrage away sectoral wage differences. Appendix Figure A.E.1 shows model-implied wage growth for the whole observed range of mortality rates and different labour shares.

Evidence from Population, Wages, and Interest Rates

In data by Buringh (2021), I show that plagued cities' populations had recovered by 1750. Additionally, I propose the number of captains reporting to live in a region as a proxy providing annual data.[31] Appendix Figure A.E.2 shows an event study on the number of captains, where the average point estimate corresponds to a 50% population drop. The number of captains climbs back to pre-plague levels after 15 years. This likely overstates both the population shock and recovery, as captains were highly skilled and mobile. Thus, labour scarcity is present in plagued regions for 15-50 years.[32]

Model-based wage growth is low compared to the Black Death. Jedwab et al. (2022) argue that high mortality disrupted trade and increased transaction costs, decreasing real wages in England and Spain for two years. Following Alfani & Murphy (2017), I suggest that markets did not unravel after the Great Northern War plague outbreak as institutions were better prepared for a mortality shock. Thus, wage increases after this plague outbreak are likely much closer to theoretical predictions.

There is evidence from the Black Death that wage increases were real. Broadberry et al. (2014) show that dairy and meat provided a larger share of calories in England after the plague.

[30] Mortality estimates in Appendix Table A.B.2 pertain to cities. 20% is in the middle between the upper and lower bound for regions. Mortality rates were lower in hinterlands (Voigtländer & Voth (2012)). Assuming 36% to hold for regional population is thus an upper bound. Bairoch et al. (1988) finds a European urbanisation rate of 11.4% in 1700, similar to the 11.9% in de Vries (2013). As a lower bound, if there had been no mortality in the hinterland, regional mortality would lie at 4.3%.

[31] This variable is the home port as in Marczinek et al. (2025). To count captains once, I omit observations in which first name, last name, and region are duplicates in a given year.

[32] I explored the Hyde data. However, these are either imputed or at the national level, using linear interpolations between centuries, without providing genuine localised annual population data. As I need regionally granular and ideally annual population data, the Hyde data are not the ideal source either.

Galofré-Vilà et al. (2018) document an average height increase of 6.59 cm between 1348 and 1400. Pfister (2017) provides evidence closer to my setting: real wages of unskilled urban male labourers in Germany doubled as the Thirty Years' War induced labour scarcity through war and plagues.

Importantly, wages rose relative to return rates on capital. Data on city-level interest rates in the Baltic and North Sea region are scarce in the early 18th century. Even the most extensive collections of sources (Huang et al. (2019), Schmelzing (2019)) cover only three of the plagued cities: Hamburg and Danzig, yet only before the plague, and Bremen, which had a mortality rate of 0.7%, whose factor proportions were thus barely affected, and which accordingly saw virtually no variation in interest rates. Thus, I instead follow Keller et al. (2021) who infer interest rates from monthly variation in grain prices. I use price data by Allen & Unger (2018) for Danzig in 1704-1793. Appendix Table A.E.1 regresses interest rates recovered from wheat, rye, and barley prices on a post-plague dummy and finds that interest rates do not increase after the plague. I conclude that both wages rise absolutely and relatively.

Step #2: Production Becomes More Capital-Intensive

Labour scarcity is followed by an increase in capital intensity along two margins. First, a higher capital-to-labour ratio is used:

$$\frac{K_{ik}}{L_{ik}} = \frac{1-\gamma_k}{\gamma_k r} w_{ik}.$$

Second, a shift into capital-intensive sectors occurs. Taking ratios of gravity equation 1.4.8 for manufacturing and plugging in equations 1.5.2 and factor market clearing yields:

$$\frac{X_{ijCM}}{X_{ijLM}} = \xi w_{iM}^{-\theta(\gamma_{CM}-\gamma_{LM})} w_{iM}^{\beta_{CM}-\beta_{LM}} \frac{CMA_{jLM}}{CMA_{jCM}}.^{33} \tag{1.5.3}$$

Three channels operate. First, capital-intensive sectors expand as they have smaller labour shares. This is the **labour cost channel**. Second, capital-intensive sectors expand as capital deepening increases productivities. For $\beta_{CM} > \beta_{LM}$, the increased capital-to-labour ratio increases productivity in capital-intensive manufacturing more than in labour-intensive manufacturing. This results in the second term in equation 1.5.3. This **productivity channel** further increases the expansion of capital-intensive sectors.

In general equilibrium, $\frac{CMA_{jLM}}{CMA_{jCM}}$ decreases, counteracting some of the shift into capital-

[33] $\xi = \frac{\alpha_{CM}}{\alpha_{LM}} \frac{\chi_{CM}}{\chi_{LM}} \left(\frac{d_{ijCM}}{d_{ijLM}}\right)^{-\theta}$

intensive exports. In a labour-intensive sector, wage growth weighs more heavily on an origin's contribution to a destination's market access. Beyond that, differences in β may lead to slower productivity growth in labour-intensive sectors. Both work to decrease $\frac{CMA_{jLM}}{CMA_{jCM}}$. In other words, plagued regions' labour scarcity and productivity advances increase the relative (export) prices of labour-intensive compared to capital-intensive goods. This relative price channel reduces the value of capital-intensive exports compared to labour-intensive exports. I control for this channel by estimating equation 1.3.2 with destination-time fixed effects.

Evidence from Capital Stocks

Combining market clearing conditions and sectoral wages 1.4.7 reveals:

$$K_{ik} = \frac{(1-\gamma_k)\zeta_k}{\gamma_k r}\Big(A_{ik}FMA_{ik}\Big)^{\frac{1}{1+\gamma_k\theta}} L_{ik}^{\frac{\gamma_k\theta}{1+\gamma_k\theta}}. \tag{1.5.4}$$

Equation 1.5.4 shows the elasticity of capital stocks with respect to employment, which is less than 1. Thus, the capital-to-labour ratio increases. While no granular data are available, I employ a proxy for capital stocks: the number of ships owned in a region.[34] Ships are an expensive infrastructure investment and highly mobile.[35] The plague leads to a larger number of ships (Appendix Table A.E.3). For two decades after the plague, significantly less ships are registered in plagued regions (Appendix Figure A.E.3). A comparison with the number of captains (Appendix Figure A.E.2) fits an elasticity of less than 1. 50 years after the plague, there were significantly more ships registered in previously plagued regions; this higher capital stock remained in place until the 1850s. The drop and recovery in population and lasting productivity gains, following equation 1.5.4, can explain these dynamics. Beyond ships, I find that plagued ports import larger volumes of capital goods (metals, tools, nails, building materials), also relative to an overall growth in imports (Appendix Figure A.E.4). This suggests that capital stocks are being accumulated.

Evidence from Production Data

In trade data, the shift *across* sectors is observable (Figure 1.3.1). I find further evidence in farm-level data from Sweden: farms shift out of labour-intensive agriculture and into capital-intensive agriculture.[36] I draw on data by Olsson & Svensson (2017). After the plague, Scania's

[34]The labour cost and productivity channels determine the allocation of capital stocks across sectors.
[35]The export expansion is not mechanically linked to an accumulation of ships. Merchants from other regions could sail to productive locations.
[36]The same pattern was observed in English agriculture after the Black Death (Clark (2016)).

farms reduce their production of labour-intensive rye and barley and shift into capital-intensive calves (Appendix Figure A.E.5).[37]

Thus, insights from trade data carry over to production data. To test the link with labour scarcity, I construct for each farm i a measure of distance to Ystad and Malmö:

$$plaguedist_{it} = \Sigma_j Distance_{ij} * Plague_{jt}. \tag{1.5.5}$$

Farms located closer to plagued Ystad or Malmö have lower values of $plaguedist_{it}$. I expect greater proximity to plagued cities to be associated with higher mortality. Accordingly, the shift into capital-intensive agriculture should be increasing in proximity to plagued cities. In Appendix Table A.E.4, I find that proximity to plagued cities is associated with higher growth in calves production and lower growth in rye and barley production.[38]

As for the shift *within* sectors, I use data for 2,344 Swedish parishes in 1690 and 1750 on land usage, animals, and population (Andersson Palm (2025)). Plagued parishes use significantly more horses, controlling for population, after the plague (Appendix Table A.E.5). This holds both in 1750 and in 1810.[39] By the 1700s, horses had become the main draught animal in much of Europe (Langdon (1986)), though oxen may still have played a role in Sweden. Thus, more horses per capita reflect capital deepening *within* agriculture, as also argued by Olsson & Svensson (2008).

In Appendix Table A.E.6, I introduce interactions with Scanian farms' soil quality. Despite the labour scarcity-induced shift out of rye, I find that farms with bad soil close to plagued cities shift into rye - the hardiest grain. Faced with both scarce labour and poor soil, farms shift into what is easiest to grow. Labour scarcity hits barley production uniformly independent of soil quality. The overall shift out of rye and into foals is driven by farms with good soil close to plagued cities. This is capital deepening in arable agriculture: horsepower is being substituted for lost labour, and particularly so with rich soil. Soil quality does not interact with other animals, supporting that this is about draught animals specifically.[40] By 1750, population had recovered in plagued parishes (Appendix Table A.B.6). Thus, both capital levels and the capital-to-labour ratio were significantly higher in 1750 and in 1810 in plagued parishes despite population recovery.

[37] While calves grow up to be cows and will produce meat and milk, some regard raising foals and calves as a metric of capital deepening in agriculture (Olsson & Svensson (2008)).

[38] The overall regional trend out of rye and barley and into calves is accounted for by time trends.

[39] The number of ships as a proxy for ports' capital stocks is also significantly higher at and beyond this time horizon (Appendix Figure A.E.3).

[40] Beyond draught animals, there is evidence of capital deepening through metal ploughs, early threshing machines and modern farm buildings, where plagued Scania stands out as an early adopter (Magnusson (2007)).

Evidence from Exports of Novel and Innovative Goods

Investment and innovation are difficult to study in this period. As a proxy, I investigate novel export goods that indicate innovation and investment into new production lines. An export expansion may indicate both more intensive usage of existing production knowledge and a creation of new production knowledge, which one may refer to as innovation. A higher capital-to-labour ratio may have allowed plagued regions to make the investments necessary to acquire new production knowledge. This link between innovation and labour scarcity finds empirical support in Andersson et al. (2020). I partial out goods that regions first exported between 1689, 20 years before the first year of the plague, and 1732, 20 years after the end of the plague.

Appendix Figure A.E.6 shows that the number of new export goods increases significantly faster in plagued regions. Novel export goods contribute a quarter to the overall extensive margin expansion. Appendix Figure A.C.18c shows that novel export goods are significantly more frequent in capital-intensive sectors. Results at the goods level are shown in Appendix Figure A.C.19. Appendix Tables A.C.7 and A.C.8 show PPML results. Even after 20 years, plagued regions continue to innovate more than non-plagued ones. This fits with the findings of Juhász, Squicciarini & Voigtländer (2024).

Durability of Capital

I argue that my finding of capital deepening after the plague is not driven by different durabilities of capital. As blast furnaces last longer than a plough, capital deepening in capital-intensive sectors may be longer-lived than in labour-intensive sectors. I provide three arguments against this. First, the time horizon over which I find capital deepening suggests continued reinvestment rather than depreciation of a one-off increase in capital stocks. Following Allen (2009*b*), suppose the aggregate depreciation rate sits at 2.4%. This implies a half-life of 29 years for the capital stock. However, the capital intensity of exports and the number of horses per capita were higher almost a century after the plague (Appendix Figure A.C.6 and Appendix Table A.E.5). Second, suppose a conservative depreciation rate of 2% in capital-intensive and a higher rate of 7% in labour-intensive sectors. The half-lives are then 35 and 10 years, respectively. Thus, long-lasting capital deepening cannot be explained by durability alone and requires continued reinvestment. Third, comparing durabilities in agriculture illustrates the limited explanatory power of depreciation. While horses and oxen may live longer than a plough lasts, their lifespans are unlikely to exceed 20 years; crucially, their working life spans are estimated at 5-7 years by Langdon (1982). Thus, the relevant depreciation rate for draught animals should be very similar

to that of tools in arable agriculture. Despite similar depreciation rates, I find strong evidence for long-lasting capital deepening in agriculture. Thus, I suggest that differences in the durability of capital do not drive my results.

Step #3: Productivity Grows; More So In High β_k Sectors

I make two assumptions to recover productivity growth (Appendix A.D.1). First, all sectors have the same labour share.[41] Second, the labour mobility friction does not vary within but only across areas, such that $\phi_{ikt} = \phi_{jkt} \forall (k,t)$ and i,j within the same area.[42]

Sectors may have different elasticities between capital deepening and productivity growth. As discussed above, I do not endogenise productivity but propose a reduced-form relationship that determines productivity as a function of factor proportions, where other determinants are absorbed by fixed effects:

$$A_{ik} = \Big(\frac{K_{ik}}{L_{ik}}\Big)^{\beta_k}.$$

If all sectors operate the same production function, capital deepening would still produce different productivity gains if $\beta_k \neq \beta_{k'}$. I expect this efficiency to be higher in capital-intensive than in labour-intensive sectors.[43] At the time, few machines and technologies existed in labour-intensive sectors that could have led to capital deepening increasing productivity growth (Atack et al. (2019), Gallardo & Sauer (2018), Coleman (1956)). To test the impact of the plague on sectoral productivity growth, I estimate equation 1.5.6 on data between 1668 and 1749:

$$\text{productivity growth}_{ikt} = \delta_k plague_{it} + \alpha_{ik} + \alpha_{akt} + \epsilon_{ikt}, \tag{1.5.6}$$

where $\text{productivity growth}_{ikt} = \delta_{ikt}$ are the fixed effects recovered from equation 1.4.9.[44] α_{ik} are region-sector fixed effects, absorbing time-invariant region-sector-specific determinants of productivity, such as availability of resources or time-invariant institutions. α_{akt} are area-sector-time fixed effects, absorbing area-wide time variation in sectoral productivity. These absorb sectoral productivity growth across an entire area, such that I compare the plague's effects of productivity growth within an area.[45] Examples of this are area- and country-wide economic

[41] Capital-intensity then comes down to different efficiencies β_k with which sectors turn capital deepening into productivity growth. Below, I put bounds on the bias induced by a violation of this assumption.

[42] This explicitly permits the reintroduction of serfom in Denmark in 1733 and worsening serfdom in Estonia (Raster (2023)). Alternative assumptions are discussed in Appendix A.D.1.

[43] If the elasticity of substitution between capital and labour were to differ sectorally, this would also show up in β_k. In the proposed model, these elasticities are 1 for all sectors under Cobb-Douglas; given the available data, I do not think that I can convincingly estimate these elasticities.

[44] See Appendix A.D.1 for details.

[45] Areas are displayed in Appendix Figure A.B.1. As before, these are equivalent to NUTS2 for most of

and trade policies. Crucially, if there is time variation in the serfdom-induced mobility friction, these are absorbed by α_{akt} as long as they do not vary within areas. The fixed effects α_{ik} and α_{akt} thus absorb the time-invariant and area-level time-varying parts of vector x_{ik} in the general functional form for productivities in equation 1.5.1. ϵ_{ikt} are the error terms. δ_k is the coefficient on the plague.

Table 1.5.1 shows significant productivity growth in capital-intensive sectors relative to labour-intensive sectors after the plague. This difference of productivity growth by factor intensity holds both within agriculture and manufacturing.

Table 1.5.1: Impact of plague on sectoral productivity growth, by sector and factor intensity

	Labour-intensive ag.		Capital-intensive ag.		Labour-intensive man.		Capital-intensive man.	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Plague Dummy	0.047		0.624***		-1.239***		0.073	
	(0.178)		(0.208)		(0.164)		(0.169)	
Mortality Rate		0.215		1.811***		-3.006***		0.635
		(0.464)		(0.544)		(0.428)		(0.441)
Fixed Effects:								
– Region	✓	✓	✓	✓	✓	✓	✓	✓
– Area x Year	✓	✓	✓	✓	✓	✓	✓	✓
Observations	50,010	50,010	50,020	50,020	50,020	50,020	50,020	50,020

Notes: Standard errors in parentheses. * p <0.10, ** p <0.05, *** p <0.01. The dependent variable is log sectoral productivity growth. The independent variable is first a plague dummy that equals 1 for plagued regions after the plague hit. The second independent variable is the mortality rate, which for half of regions is imputed as the predicted value from regression results presented in Appendix Table A.B.4.

Sectorally differing productivity effects come down only to differences in β_k under the assumption that all sectors have the same labour share. A violation of this assumption does not affect the level effect for productivities, which is absorbed here and for which I find significant estimates below. Wages can only rise after the plague: as exports go up, there needs to be productivity growth in a Ricardian framework. However, unaccounted for differences in labour shares lead to an overestimation in sectoral differences in β_k. In Appendix Figure A.E.7, however, I show that for model-implied wage growth and even extreme differences in labour shares, this scaling factor leads to an overestimation in point estimates of only 17.3-37.6%. Further, I find similar results when accounting for spatial correlation (Conley (1999)) in Appendix Figure A.E.8.

How did productivities respond once the adverse labour supply shock had been overcome (Appendix Tables A.B.5 and A.B.6)? Extending the productivity analysis to the long-run, I find immediate productivity growth for both agriculture and manufacturing in the first post-plague decade, with capital-intensive manufacturing (agriculture) seeing faster productivity growth than

Europe; for the UK, Germany, France, and Spain, they are NUTS1, to have similarly sized areas and avoid areas with a single port (Appendix A.B).

labour-intensive manufacturing (agriculture). Plague-induced productivity changes persist for half a century longer than the underlying changes in factor proportions. This suggests that temporary shocks to factor proportions can produce persistent productivity changes.

Evidence from Production Data

To complement the gravity-based recovery of productivity, I estimate a Cobb-Douglas production function and recover logarithmic productivities (TFP) as residuals. I use agricultural data by Andersson Palm (2025) for Swedish parishes in 1690 and 1750. As revenue for labour-(capital-)intensive agriculture, I use the value of harvests (animal output[46]). I then regress logarithmic revenue on logarithmic population (labour), sowed acreage (land), and draught animals (capital), and recover the residuals. Table 1.5.2 shows that TFP levels are significantly higher in plagued parishes. The effect is larger in capital-intensive than in labour-intensive agriculture. All this is in keeping with the previously-established productivity findings that relied on recovery from the gravity equation.

Table 1.5.2: Parish-level TFP growth after the plague

	Agricultural TFP by Factor Intensity					
	(1)	(2)	(3)	(4)	(5)	(6)
Post Plague	0.716***	0.749***	0.693***	0.736***	0.729***	0.748***
	(0.0463)	(0.0520)	(0.0467)	(0.0559)	(0.0485)	(0.0528)
Post Plague x Capital Intensive	0.399***	0.322***	0.304***	0.142	0.274***	0.715***
	(0.0828)	(0.0841)	(0.0842)	(0.0886)	(0.0906)	(0.116)
Outcome	Log TFP	TFP	Log TFP	TFP	Log TFP	TFP
Draught Animals	Horses and Oxen	Horses and Oxen	Horses	Horses	Oxen	Oxen
Fixed Effects:						
– Parish x Factor Intensity	✓	✓	✓	✓	✓	✓
– Factor Intensity x Year	✓	✓	✓	✓	✓	✓
Observations	3,914	3,914	3,748	3,748	2,632	2,632

Notes: Standard errors clustered at the parish level in parentheses. * p <0.10, ** p <0.05, *** p <0.01. The dependent variables are parish-level (logarithmic) TFP by factor intensity estimated from regressing logarithmic harvests or stock of animals (for labour- and capital-intensive agriculture, respectively) on logarithmic population, sowed acreage, and horses and oxen (columns 1-2, or only horses in columns 3-4, or only oxen in columns 5-6) as draught animals. The independent variable is a post-plague dummy and its interaction with a dummy for capital-intensive agriculture.

Recovering productivity as a residual from regressing revenues on inputs has well-documented issues: beyond simultaneity bias, the residual mixes TFP with demand-side price variation (Klette & Griliches (1996)). In this period, however, there is no information on intermediate inputs (Levinsohn & Petrin (2003)) or investments (Olley & Pakes (1996)). As the data are recorded every 60 years, the scarcity of time periods thwarts me from employing estimators such as Ackerberg et al. (2015). Thus, recovering TFP as residuals may be the most suitable

[46]Specifically the number of calves, cows, sheep, goats, pigs, and heifers; the output of e.g. milk and meat should be proportional to this.

approach. I follow a literature estimating firm-level production functions (Griliches & Mairesse (1984), Griliches & Mairesse (1998), Hall & Mairesse (1995)). These papers use revenue, capital, and employment as these are the only variables in typical firm accounts.

Similarly, the residual approach faces the objection of input mismeasurement: after the plague, the quality of labour may have risen, as survivors may be systematically healthier. Capital utilisation may have changed, too, and land quality per worker may have improved. Such factor quality upgrading contaminates the residual which I claim to capture TFP. However, I argue that the larger TFP growth in capital-intensive sectors is hard to explain by factor quality upgrading, which should be uniform across sectors. Further, the post-plague observation is in 1750, four decades after the plague; selection effects in surviving workers should have disappeared by then. Below, I also rule out differential skill acquisition that may have improved human capital. This suggests that, indeed, the plague induces TFP growth.

There is also historical evidence for productivity growth after the plague. Clark (2016) and Broadberry et al. (2014) show that the Black Death was followed by productivity gains. For this outbreak, Appendix Figure A.E.9 shows a brief drop in output per farm, adjusting for size, based on Scanian data (Olsson & Svensson (2017)). Production surpasses pre-plague levels after a few years, consistent with a productivity effect.

Beyond agriculture, one Swedish manufacturing industry that saw important changes was ship building. Using data on ships from the Swedish East India company, I show in Appendix Figure A.E.10 that the vast majority of ship yards built in Sweden after the plague was built in previously plagued regions.[47] These also built the largest and most productive ships, as measured by the number of successful trips to East Asia.

1.5.1 Evidence from Factor Frictions

The mechanism can be validated by studying settings where it is constrained by factor market frictions. I study two such frictions: for labour, serfdom which limits the mobility between cities and their hinterlands. For capital, I investigate pre-war variation in the depth of capital markets which should affect the degree of capital deepening.

[47] Construction of the Polhemsdockan in Karlskrona began in 1717, five years after the plague hit the city. This dry dock used pumps rather than tides and is considered by some the Eighth Wonder of the World.

Labour Frictions

Serfdom severely limited labour mobility between cities and their hinterlands. I analyse how capital deepening and productivity growth differ for regions with serfdom (Section 1.2.3 and Appendix Figure A.E.11). Serfdom hindered rural-to-urban labour mobility by binding serfs to their lords' lands. Further, serfs in the Eastern Baltic were not allowed to learn a trade and typically spoke Estonian or Latvian. Skill and language barriers thus further prevented serfs from moving to (mostly German-speaking) cities.

Under full labour mobility, wages are equalised between a city and her hinterland. However, under mobility restrictions, higher urban wages are no longer equalised to those in the hinterland as peasants cannot move. Denmark presents a testing ground for this hypothesis as serfdom was re-introduced in 1733 (Gary et al. (2022)). There was no wage gap between urban and rural workers before 1733; however, a gap appears just after serfdom was brought back (Appendix Figure A.E.12). I formalise this wedge of ϕ_i between urban and rural wages: $w_{iM} = (1 + \phi_i)w_{iA}$.

I consider two forms of labour scarcity. Under labour mobility, labour scarcity from higher urban mortality rates (Voigtländer & Voth (2012)) is equalised across sectors. I refer to this as *absolute* labour scarcity, as for all sectors, labour has become scarce. With serfdom, however, labour mobility does not equalise labour scarcity across sectors. Beyond absolute labour scarcity, cities suffer from *relative* labour scarcity as workers from the hinterland cannot move to benefit from higher urban wages. Thus, labour is artificially scarce in cities and artificially abundant in rural areas. This *relative* labour scarcity manifests itself in a wage wedge.

In Appendix A.E.3, I show how this logic affects production and exports in my model. Urban manufacturing wages increase more under serfdom; thus, the shift into a higher capital intensity of exports should be larger in plagued regions with serfdom. In agriculture, the opposite is true: under serfdom, rural agricultural wages rise less after the plague, leading to a smaller shift into capital-intensive agricultural exports under serfdom. Appendix Table A.E.8 confirms these expectations.

Capital Frictions

In imperfect capital markets, capital deepening is constrained by financial development: ports with deeper pre-war capital markets should undergo stronger capital deepening. Data on loans and interest rates are very limited for this period. I summarise three tiers of financial development for large ports in the area: Hamburg, Amsterdam, and Bremen were the most developed (Huang et al. (2019), Carlos & Neal (2011), Denzel (2010)). Danzig may be grouped in a second tier

(van der Wee (1977)). Stockholm, Königsberg, Riga, Reval, and Lübeck were the least financially developed major ports.[48] Going deeper, empirical evidence for financial variables is lacking for virtually all ports.

Given these constraints, I borrow an insight from Böhm-Bawerk (1889) and Antràs & Kulesza (2026): financially developed regions have a comparative advantage in sectors with a high average period of production (APP). In sectors where more time passes between spending on inputs and revenue raised from selling final goods, financing costs are higher. Thus, ports' pre-plague APP in exports reflects their financial development.

I construct each port's export-weighted pre-war APP.[49] I ask whether the shift into a higher capital intensity and the export expansion are stronger for ports with a higher pre-war APP. Throughout, I control for all interactions with pre-war export levels by sector, as APP is higher in manufacturing and export-weighted APP is mechanically higher in larger ports. In Appendix Table A.E.10, I find that, even when controlling for export volumes and sectoral structure, plagued ports with a higher pre-war APP see a significantly larger shift into capital-intensive exports and a significantly larger export expansion. Appendix Table A.E.11 shows that manufacturing productivity growth is stronger in regions with a higher pre-war APP. This suggest that the observed effects are explained by capital deepening: my findings are significantly stronger for ports that were less financially constrained and could more easily engage in capital deepening.

1.6 Alternative Mechanisms

As an alternative mechanism, I can rule out pure Heckscher-Ohlin forces and show that Ricardian forces account for more than 90% of variation in a hybrid framework. I present historical evidence against the presence of a factor bias in technical change. However, I cannot rule out that the true model is a hybrid containing Ricardian and Heckscher-Ohlin forces; similarly, I cannot rule out that technology is factor-biased. I nonetheless argue that productivity growth after the plague is required to explain my findings. Further, I find no evidence for human and social capital deepening after the plague. I also conclude that "venting out" and market power are ill-suited to explain my findings. Finally, I test for the presence of non-homotheticity (Fieler (2011)).

[48] For example, shadow interest rates in Stockholm were twice as high as in Bremen (Huang et al. (2019), Edvinsson (2011)), indicating poor access to credit. Königsberg, Riga, and Reval do not even appear in financial series before 1700 (Denzel (2010)), suggesting that credit may have depended entirely on the financial markets of Hamburg and Danzig.

[49] All details in Appendix A.E.3. Ranking ports by their average APP, I find that the major ports are ranked as follows: Hamburg, Bremen, Amsterdam, Reval, Riga, Danzig, Stockholm, Königsberg, Lübeck. This matches the above-described historical evidence on financial development very well.

Heckscher-Ohlin Forces

I relate my model to Romalis (2004), Morrow (2010), and Fadinger & Fleiss (2011) and argue that Heckscher-Ohlin forces may be present, but are neither required nor sufficient to explain the findings of this paper; in a hybrid model, more than 90% of the difference in sectors' supply-side responses are driven by Ricardian forces, leaving only a small role for factor forces.

Fadinger & Fleiss (2011) propose a version of their model in the spirit of Eaton & Kortum (2002) that does not require production data. In the resulting gravity equation[50] exporter-industry fixed effects capture TFP but are contaminated with the interaction of factor abundance and factor intensity:

$$\hat{\delta}_{ikt} \propto \frac{\exp(\lambda_{ikt}/\beta_k)}{\prod_f (w_{it}^f)^{a_{fk}/\beta_k}} = \frac{\exp(\lambda_{ikt}/\theta_k)}{(w_{it})^{\gamma_k/\theta_k} r^{(1-\gamma_k)/\theta_k}}, \tag{1.6.1}$$

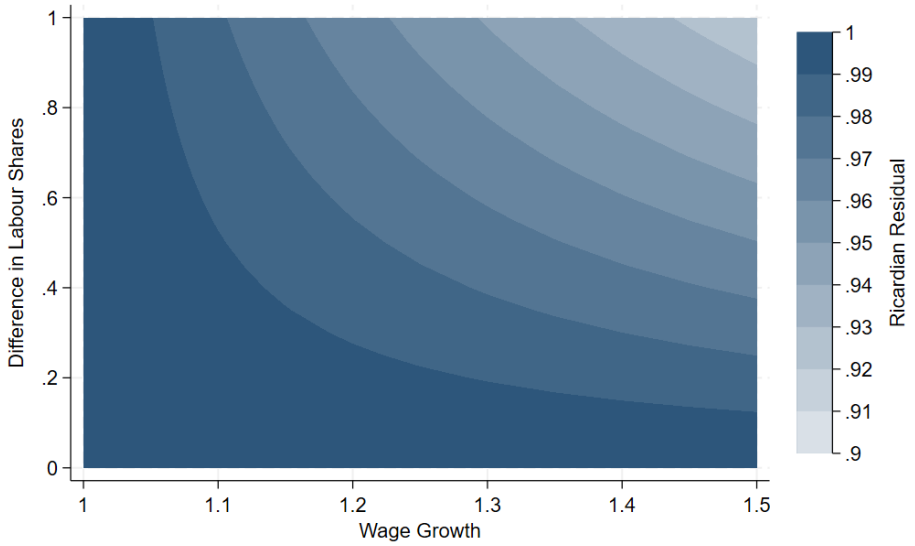
where $\exp(\lambda_{ikt}/\beta_k)$ is productivity and w_{it}^f are factor prices and a_{fk} are factor shares. The identity introduces the notation used in my paper, where $\beta_k = \theta_k$, $w_{it}^L = w_{it}$, $w_{it}^K = r$, $\alpha_{Lk} = \gamma_k$, and $\alpha_{Kk} = 1 - \gamma_k$. While I cannot precisely separate Ricardian forces from factor terms without data on factors' prices and shares in the production function, I calculate the scaling factor for different wage growth and labour share figures. In Figure 1.6.1, I show the extent to which my productivity growth estimates need to be scaled down. For model-implied wage growth of 6.6% and a difference in labour shares of 0.5, the Ricardian term needs to be scaled down by 0.64%[51]; even for large wage growth and labour share differences, Ricardian forces continue to account for more than 90%.[52]

[50] Equation (A.5) of the Online Appendix in levels: $M_{ijk} = \frac{\exp(\lambda_{ik}/\beta_k)}{\prod_f (w_i^f)^{a_{fk}/\beta_k}} \cdot \frac{1}{\tau_{ijk}^{1/\beta_k}}$

[51] As to the (here unidentified) level effect of productivity growth, the estimates of δ_{ikt} would actually need to be scaled *up* by 0.6% in a hybrid framework.

[52] For $\theta = 2$ as a very low estimate of the trade elasticity, this figure is still above 80% throughout.

Figure 1.6.1: Scaling factor required to turn δ_{ikt} into pure Ricardian term



Notes: This figure shows the scaling factor by which recovered δ_{ikt} would need to be divided to clean for Heckscher-Ohlin forces in the hybrid model of Fadinger & Fleiss (2011) and to arrive at the remaining Ricardian forces. When taking ratios of equation 1.6.1 with a reference period and sector, the scaling factor is $\left(\frac{w_{it}}{w_{it'}}\right)^{\frac{\gamma_{k'}-\gamma_k}{\theta}}$.

Outside of this hybrid framework, could pure Heckscher-Ohlin forces rationalise my findings? In Table 1.5.1, I show that labour-scarce exporters see their exporter-industry fixed effect drop relatively in labour-intensive sectors. This is consistent with two mechanisms: first, Heckscher-Ohlin forces are operating, and TFP changes may be absent or uniform across sectors. In the latter case, the symmetric increase in TFP is absorbed by fixed effects. The observed rotation into capital-intensive exports is then truly the Rybczynski effect. Second, there is sector-biased TFP growth favouring capital-intensive sectors, and Heckscher-Ohlin forces are not necessary.

I first present arguments against the first explanation. The export share expansion in Figure 1.3.2 cannot be purely reflective of Heckscher-Ohlin forces. Combining Romalis (2004) with a national accounting logic, reduced labour supply increases wages, and under standard assumptions, GDP falls. Imports fall and domestic absorption likely falls, too. Under balanced trade, overall exports must then also fall. For exports to, instead, increase, domestic absorption would have to fall by more than income. Contradicting this point, I find in Appendix Figure A.C.20 that imports rise after the plague. Additionally, Appendix Figures A.C.10 and A.C.12 show that the export expansion is concentrated in capital-intensive sectors. The absorption story cannot account for this pattern. In the absence of falling domestic absorption, wage increases reduce competitiveness across sectors, just less strongly in capital-intensive ones. Exports, then, cannot increase in the absence of TFP growth that shifts up competitiveness. Further, the finding of productivity growth in levels that is stronger in capital-intensive sectors (Table 1.5.2) directly

contradicts the prediction of a Heckscher-Ohlin framework that output patterns are shifted, yet that productivity residuals are unaltered. This shift in TFP is direct evidence of Ricardian forces at play.

Thus, there is no support for the pure Heckscher-Ohlin interpretation. I propose that TFP growth favouring capital-intensive sectors is in the driver's seat. For exports to rise in Costinot et al. (2011), controlling for bilateral and demand-side determinants, wages need to drop or productivity needs to rise. The export expansion is thus indicative of productivity growth that is large enough to dominate plague-induced wage growth. Further, I argue below that a pure Heckscher-Ohlin framework mechanically implies changes in relative imports. Instead, I find that relative imports are unaffected by the plague, which a Ricardian framework can rationalise.

However, both Ricardian and Heckscher-Ohlin forces may be operating. Under the proposed reduced-form dependence of productivities on capital deepening (equation 1.5.2), factor proportions enter through the backdoor, and are proposed to determine fundamental productivities. The difference is that under Heckscher-Ohlin, the production frontier is held fixed and factor are reallocated across sectors. My story is that the production frontier itself shifts as a result of capital deepening, which explains the rising capital intensity of exports, growing market shares in destination ports, and TFP growth. The above decomposition in the framework of Fadinger & Fleiss (2011) suggests that indeed Ricardian forces account for the vast majority of my findings.

Factor-biased technological progress

Labour scarcity could lead to directed (e.g. labour-saving) technical change and productivity growth (Alfaro et al. (2024), Acemoglu (2010), Acemoglu (2002)). This is consistent with the proposed mechanism. However, history does not support the presence of a labour-saving factor bias in this era. Labour-saving mechanisation in crop agriculture began succeeding only in the 19th century (Gallardo & Sauer (2018)). An early example of a crop harvesting machine that explicitly substituted labour for capital was the cotton gin by American inventor Eli Whitney, patented only in 1794. Focusing on the 17th century in England, Coleman (1956) argues that labour-saving technologies in agriculture were practically non-existent. Those that existed, such as mills, spinning wheels, and water-powered blast furnaces, were centuries old. In manufacturing, Atack et al. (2019) document how innovative machinery consolidated multiple hand tasks into a single machine task. They consider this the most labour-saving type of technological transition, which, however, only came to bear during the 19th century. Coleman (1956) argues for a political will to absorb rather than substitute labour in this period, resulting in opposition to

the development of labour-saving devices in agriculture and manufacturing alike. This points to unbiased technical change.

On factor bias in international trade, Trefler (1993) reconciles factor content theories with the empirical rejection of factor price equalisation by introducing factor-biased technological progress. The threat to separating wages from productivities is that the factor bias correlates with labour shares and that I am comparing sectors with different labour shares. In that case, wage growth contaminates the productivity comparison. While I cannot test for such a correlation directly, the findings of Table 1.5.1 permit a role for factor-biased rather than Hicks-neutral productivity growth. In this case, an interpretation as pure TFP overstates the Ricardian story. Above, I study the bias introduced by a violation of my assumption on labour shares: I find that productivity growth remains largely in place as the upper limit on the inflation of productivity estimates is 32%. Thus, factor-biased technological progress is consistent with my findings, yet leaves the productivity growth story largely in place.

Human and Social Capital Deepening

Equation 1.5.2 may contain human and social capital if they have sectorally varying components. After the plague, survivors may choose to invest more in their and their offsprings' human capital; similarly, there may be selection in survivors in that better connected individuals have higher survival chances. If these adjustments have a sectoral component, human capital deepening and social capital deepening may enter in equation 1.5.2. I do not expect this to be a major concern: First, human and social capital are typically region-time variation, which is absorbed in my recovery of productivities. An increase in human or social capital across sectors is inconsequential for my productivity estimates. Second, this paper's setting in an early modern period implies that sector-specific human capital was very limited, suggesting that the absorbed level effect of any potential human capital deepening would be the main concern.

To more rigorously address these concerns, I test whether there is any evidence for human or social capital deepening. If these additional types of capital were sectorally-varying and responding to labour scarcity, I would expect to see a relative export shift into goods that use these factor intensely, applying the logic of equation 1.5.3. I construct measures of goods' skill intensity and contract intensity (see Appendix A.E.4). In Appendix Table A.E.12, I show that across specifications the only consistently positive and significant estimates are for a shift into a higher physical capital intensity. There is no shift into a higher contract or skill intensity of exports. This analysis has to be caveated by the fact that both measures of human and social

capital intensity are based on modern data for the U.S. I assume that there is a sufficiently high correlation over time in goods' relative skill and relative contract intensities. The evidence presented here suggests that it is indeed physical capital deepening, and not human or social capital deepening, that explains the post-plague export expansion.

Venting Out & Market Power

Almunia et al. (2021) propose non-constant marginal costs to explain "venting out", for which this historical setting provides an ideal testing ground: while the Black Death hit all of Europe, the 1708-1714 plague outbreak left most of Europe as unperturbed export markets to counteract the local demand shock of the plague. Under venting out, regions would increase their exports in line with my findings. I suggest the time dimension as a crucial difference, however. Almunia et al. (2021) track exports for four years, while this paper studies trade for decades after the plague. Short-run marginal costs are unlikely to explain this persistent export boom.

Market power may reconcile short-run marginal cost changes with my long-run findings. If plagued regions capture market shares through lower marginal costs, a combination of market power and switching costs could explain persistent market share gains. However, I present evidence in Appendix Table A.E.15 that export prices of plagued origins drop significantly. While in the short-run this may be explained through venting out, export prices of capital-intensive goods keep falling for 30 years. This pattern is incompatible with increased market power, but fits with productivity growth.

Non-Homotheticity

Under non-homothetic demand, demand shifts relatively into more income-elastic goods after labour scarcity leads to real wage gains. Voigtländer & Voth (2012) place non-homothetic demand at the centre of their argument. The authors study the much earlier Black Death of 1346, which encompassed all of Europe. Therefore, my paper is not a test of Voigtländer & Voth (2012). However, as non-homothetic demand is frequently discussed in this context, I gather evidence that it played no role for this outbreak.[53] I assume that capital-intensive agriculture (manufacturing) is more income-elastic than labour-intensive agriculture (manufacturing).

In Appendix A.E.5, I show how my model nests non-homothetic demand (Fieler (2011)) and present three pieces of evidence against its relevance here. First, I show that my estimates of

[53]The only mention reminiscent of non-homothetic demand in the historical literature for this plague outbreak is discussed in Frandsen (2009): in some places, it was believed that textiles and bedding spread the disease, which led to survivors burning the textiles of those who died. After the plague, this might imply a relative demand shift into textiles, which are summarised as labour-intensive manufacturing here. I show below that there is no indication for a relative demand shift into textiles.

productivity growth are robust to non-homothetic demand. There is thus no contamination of productivity growth estimates with demand-side factors. The question is therefore not whether there is one or the other but whether there is both productivity growth and non-homotheticity.

Second, I compare price predictions between my model and one of non-homothetic demand. As demand shifts relatively into more income-elastic capital-intensive goods, these should become relatively more expensive. Under the proposed productivity mechanism, however, capital-intensive goods should become relatively cheaper. In Appendix A.E.5, I show that prices for capital-intensive goods decline relatively in plagued ports.[54] Further, the Soundtoll data allow me to test granular price predictions: comparing import prices to export prices permits disentangling supply-side from demand-side factors. Under non-homotheticity, changes in relative demand translate into relative import price changes. Controlling for the supply side, I find that relative import prices of capital-intensive goods are unaffected by the importer's plague status. This suggests homothetic demand. Export prices, on the other hand, are directly affected by the proposed productivity growth: export prices of capital-intensive goods should decline relatively. This is precisely what I find.

Third, I test directly for a relative demand shift after the plague by constructing a test on relative imports (Fieler (2011)). There is no evidence of any relative demand-side changes, which also rules out that sector-specific embodied technology linked to increasing imports drives my TFP findings. Constant relative imports actually provide further support for a Ricardian rather than a Heckscher-Ohlin framework. In the latter, labour scarcity should shift exports into more capital-intensive sectors, mirrored by imports shifting into more labour-intensive sectors. However, this mirroring need not be present in a Ricardian framework with heterogeneous TFP growth, which can shift the export composition without mechanically shifting the import composition. The relative ordering of labour-intensive goods may be unaffected; further, the composition of imports is subject to the supply side's technology distribution, which was not affected by the plague. The fact that relative imports do not change accordingly supports a Ricardian framework of heterogeneous TFP growth. All three elements of this analysis thus find no support for non-homothetic demand and instead support this paper's productivity growth mechanism.

[54] These price effects reconcile rising meat and dairy consumption after the plague (Broadberry et al. (2014)) with homothetic demand.

1.7 Counterfactual

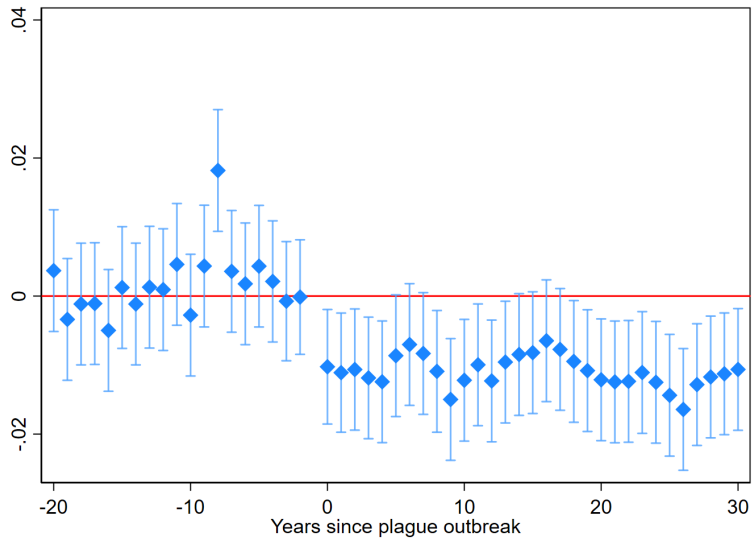
In a counterfactual analysis, I shut down productivity growth. I find that market shares of plagued regions would have contracted rather than expanded as wages rise and competitiveness falls.[55] Counterfactual market shares t years after the plague can then be decomposed into a time-invariant term, s_{ijkPRE}, that contains productivities, trade costs, and incomes, and a time-varying component reflecting employment:

$$s^c_{ijkt} = \frac{L_{ikt}}{L_{ikPRE}}^{\frac{\gamma_k \theta}{1+\gamma_k \theta}} s_{ijkPRE}. \tag{1.7.1}$$

As no annual population data are available, I suppose an upper bound for population recovery and impose exponential population growth. In Figure 1.7.1, I repeat the event study in equation 1.3.2 but replace post-plague export shares of plagued regions by counterfactual shares. I find significantly lower export shares as a smaller population led to higher wages and lower competitiveness. This stands in stark contrast to my empirical findings. To get from the average point estimate in counterfactual export shares (a contraction of 1.03 p.p.) to the observed expansion of export shares (by 0.98 p.p.), average productivity needs to rise by 210%. This - admittedly high - estimate mirrors the magnitudes found in the TFP estimation in Table 1.5.2, where the point estimate for labour-intensive (capital-intensive) TFP growth (column 1) implies TFP growth of 105% (205%).

[55] I plug in gravity equation 1.4.8, equation 1.4.7 for wages and equations 1.4.5 for firm and consumer market access to arrive at an equation defining market shares as a function of productivities, sectoral employment, trade costs, wages, and total income. A_{ik} is fixed. All details are in Appendix A.F.

Figure 1.7.1: Counterfactual market shares without productivity channel



Notes: Counterfactual export shares, shutting down productivity channel. Estimate equation 1.3.2 on counterfactual export shares after the plague and observed export shares before the plague.

1.8 Conclusion

This paper studies the relationship between labour scarcity and productivity growth. I establish novel empirical facts: First, capital-intensive exports expand relatively after the plague. Second, plagued regions expand their market shares in destination markets.

To conceptualise my findings, I build a Ricardian model. I suggest that productivity is a function of factor proportions, where other determinants are absorbed by fixed effects. From this model, I back out productivity and show that productivity grows, and more so in capital-intensive sectors. I propose that capital deepening leads to productivity growth and provide additional supporting evidence from farm- and parish-level agricultural data. These productivity changes are in place for almost a century, even though populations had recovered after 40 years.

I also present arguments against alternative mechanisms. Two alternatives can rationalise my findings: a hybrid model of Ricardian and Heckscher-Ohlin forces; and factor-biased technological progress instead of TFP. Nonetheless, to explain the empirical findings of this paper, both require productivity growth, though the point estimated would differ. In counterfactual analysis, I shut down the productivity channel and find that exports decline. The difference between this counterfactual export contraction and the observed export expansion is well-explained by and fits in magnitude with my productivity growth estimates.

The findings of this paper suggest several take aways. Many industrialised economies grapple with persistently low productivity growth rates. At the same time, their societies are aging

and preferences on working hours appear to be changing. Further, long-term health conditions are reducing labour force participation rates in several large economies (Di Meo & Eryilmaz (2025)). Thus, labour scarcity in industrialised economies reflects both declines in the working age population and in labour force participation. I argue that this labour scarcity can lead to productivity growth if agents adjust to changes in factor prices and comparative advantage.

Chapter 2

Networks in Trade - Evidence from the Legacy of the Hanseatic League

This is the latest working paper version, which has now been published: *Max Marczinek, Stephan Maurer, Ferdinand Rauch, "Networks in trade — Evidence from the legacy of the Hanseatic league," Journal of International Economics, Volume 157, 2025.*[1]

Max Marczinek[2]**, Stephan Maurer**[3]**, Ferdinand Rauch**[4]

Abstract. We study trade networks following the decline of the Hanseatic League, using a novel trade data set that covers cities and captains in Northern Europe over 190 years. By the time of its dissolution in 1669, trade on former Hansa routes is within predictions from a gravity framework. However, the Hansa continue to shape the composition of trade: Trade between cities in the Hanseatic network continues to be facilitated by Hanseatic captains for centuries. Our paper highlights the long-run stability of commercial and social networks, which persist when other economic effects do not.

JEL Classification Codes: F14, N73, N93

Keywords: Hanseatic League, Network, Gravity, Trade

[1]We would like to thank the Soundtoll team for making their data readily available, and for Bernd Wurpts for making the data of Poeck (2010) available to us in electronic format. This research has been supported by a grant from the Brasenose College Research Fund.

[2]University of Oxford

[3]University of Edinburgh, UPF Barcelona School of Management, and CEP

[4]University of St Gallen

2.1 Introduction

In this paper, we investigate how trade networks develop after their formal dissolution. Trade networks can reduce frictions and transaction costs, thereby fostering connections and increasing trade and welfare (Greif (1993), Greif et al. (1994), Gomtsyan (2022)). However, the formation of networks is costly. In order to compare high set up costs to dynamic gains over time, we need to understand their long run behaviour. At the same time, investigating trade networks over a long period of time has been difficult due to demanding data requirements. On the one hand, researchers need to observe trade flows for a long time, which tends to limit the granularity at which they observe trade. On the other hand, one needs highly granular data to observe the names or networks of traders, or a reasonable proxy for it. Therefore, long-run trade studies such as O'Rourke & Williamson (2001) often speak more to unilateral determinants of trade.

We address these issues by using a historic laboratory: the demise of the Hanseatic League in the 17th century. By overcoming trade frictions and weak contract enforcement, this confederation dominated trade in the Baltic region during the Middle Ages. It ended in 1669, after a long period of gradual dissolution.

We use Danish tax register data that records seaborne transits between the North and the Baltic Sears for the years 1668 to 1857. We use these data to build a panel of city-level, high resolution trade flows over 190 years. We not only observe origin and destination of ships, but also captains' home towns and last names, which we use as proxies for the captains' network memberships. This allows us to estimate how the former Hansa network shaped trade long after its dissolution. Our paper is among the first set of papers in economics that develop and use this dataset. Alder et al. (2020) and Waldinger (2022) use the Sound Toll Data to study the impact and consequences of environmental shocks, using more aggregated versions of this dataset than we do in this paper. Gomtsyan (2022) uses the Soundtoll data to study merchant networks in cities. Vedel (2024) investigates unilateral trade responses to a change in first-nature geography using these data. We believe that our paper is the first one that uses the Soundtoll data in its richness and in a fully specified gravity model.[5] The Soundtoll data only records trade that passes the Sound near Copenhagen. Trade between cities that are both to the east or both to the west of the Sound are missing. However, we verify below using simulations that this limitation does not introduce bias into the gravity framework.

Our main hypothesis is that trade between merchants and captains that all belonged to

[5]A more recent paper by Raster (2024) uses the full Soundtoll data at the voyage level to study the role of pioneers in the creation of trade links.

the Hanseatic network, which we define as trade by Hanseatic captains on Hanseatic routes, accumulated trading capital. Trading capital, broadly defined, encompasses physical, social, and human capital lowering bilateral trade costs (Beestermöller & Rauch (2018)). Forming relationships and building trust is also an integral part of trading capital. Its gradual depreciation has been used to explain persistent trade patterns (Head et al. (2010)). During its existence, membership in the Hanseatic League brought tax and legal privileges that were linked to being a citizen of a Hansa town. This incentivized Hanseatic traders to build particularly large stocks of trading capital on their key routes, consisting of harmonized legal and economic institutions, a common language, family connections, and trust in captains. Such trading capital will have led to lower trade costs and thus more trade among the League's members. We test this hypothesis both for the network of cities and for the network of merchants and captains, incorporating a proxy for captains' identities.

We first consider trade between former Hansa cities and their key trading partners: Does the size of aggregate trade flows between two cities, measured by the number of ships sailing between them, depend on past membership in the Hanseatic League? A persistent trade effect of formerly shared institutions has for example been documented in other contexts by Head et al. (2010), Beestermöller & Rauch (2018) and Jacks et al. (2020). Using city-level gravity equations, we do not find evidence for such trade persistence. Motivated both by the historical literature (Dollinger (1964)) and the recent literature on the importance of social connections as a determinant of trade (Cristea (2011), Munshi (2014), Campante & Yanagizawa-Drott (2017), Fang et al. (2020), Bailey et al. (2021)), we then add captains' network membership as a third dimension, measured by the captain's residence or last name. We incorporate these by creating a larger dataset where an observation is a combination of origin, destination, and a captain's residence or last name. In line with our hypothesis that Hanseatic traders invested particularly strongly into accumulating trading capital within the network, we find that trade on Hanseatic-led ships between former Hanseatic cities is considerably larger until well into the 19th century, nearly two centuries after the end of the Hanseatic League. Our mean point estimate across time and different specifications is equivalent to a tariff reduction of 8.8%. Thus, membership in the network continues to shape trade flows within the network after its formal dissolution. These results are not driven by the fact that those facilitating trade between two cities (or countries) predominantly resided in one of the two cities. Our results are further robust to different codings of Hanseatic cities and to different measures of trade flows.

In terms of mechanisms, we do not find that Hansa-specific trading capital on Hanseatic

routes was reflected in a more homogenised set of units and measures, which may have reduced trade costs. We do find, however, that ships led by Hanseatic captains carried a larger variety of goods on Hanseatic routes, showcasing their stronger integration. This finding also speaks against the hypothesis of correlated consumption preferences among Hanseatic cities.

Our paper contributes at least to four literatures. First, we add to a growing body of literature on trade networks. The role of networks in trade and development has received increasing attention (Munshi (2014)), in particular due to their capacity to overcome weak institutions (Nunn (2007), Atkin & Khandelwal (2020)) and to reduce search frictions. Xu (2022) shows that networks increase welfare.[6] Our main result shows that the network status of traders and captains continues to shape trade composition for a long time. This finding amplifies the importance of networks in trade and development. However, our results also caution that network effects might not always be apparent in bilateral trade flows, and instead be more relevant for the identity of the intermediators of trade, i.e. for the "who" rather than the "how much".

Our paper also relates to a long literature on how institutions can be passed down from generation to generation (Spolaore & Wacziarg (2009), Guiso et al. (2016)). For example, Greif (1993) shows how a coalition of Maghribi traders built a reputation mechanism to reduce transaction costs in trade relationships in the 11th century. In a similar vein, Greif et al. (1994) interpret medieval merchant guilds such as the Hansa as a way of permitting effective collective action against rulers of trade centers, leading to more efficient trade. Puga & Trefler (2014) present evidence that a small group of merchants blocked access to trade in medieval Venice, using marriage alliance to build trust and monopolize trade. Our paper relates to this literature by estimating the role of networks in a high-resolution and long-period dataset. Hanseatic traders and captains remain biased towards the Hansa network long after its dissolution.

We further contribute to the literature that shows that social connectedness and cultural similarity have measurable effects on trade (Rauch (2001), Rauch & Trindade (2002)). A common language facilitates trade (Melitz (2008), Melitz & Toubal (2014)) and trust, arising from cultural similarity and a shared history, increases trade between European countries (Guiso et al. (2009)). At the same time, Guiso et al. (2016) and Nunn & Wantchekon (2011) show how (mis-) trust and social capital can persist over many generations. Burchardi et al. (2018) find that reduced information frictions, rather than taste similarities or factor endowments, drive the effect of ancestry composition on foreign direct investment sent and received by US firms.[7] Burchardi

[6] A more recent literature in trade also looks specifically at firm-to-firm networks, see for example Bernard et al. (2019), Atkin & Khandelwal (2020), and Boken et al. (2023).

[7] In the context of online labour markets, Xu (2015) finds an important role of collective reputation as proxied by nationality.

& Hassan (2013) link individuals' social connections to regional economic growth. Bailey et al. (2021) find that countries that are better connected on Facebook trade more, while Felbermayr & Toubal (2010) highlight the relationship between trade and cultural proximity as measured by voting patterns in the Eurovision song contest. Furthermore, cultural networks of migrants persist for decades (Gould (1994), Rauch & Trindade (2002), Dunlevy (2006), Parsons & Vézina (2018)).

Finally, we contribute to a literature on trade persistence. For example, centuries after the demise of the Roman Empire, regions that were better connected during Roman times still have more business links (Flückiger et al. (2021)). Similarly, inter-German trade patterns persisted after German reunification (Nitsch & Wolf (2013)), a banking crisis in the late 19th century disrupted trade patterns for several decades (Xu (2022)), and trade flows in 1870 strongly predict trade flows in modern times, even conditional on modern gravity variables (Campbell (2010)). On the other hand, when colonies become independent, their trade volumes with their former metropolis revert gradually to a gravity benchmark (Head et al. (2010), Jacks et al. (2020)). A similar decline arose in the case of former parts of the Austro-Hungarian Empire after the fall of the Iron Curtain (Beestermöller & Rauch (2018)). Overall, estimates of the degree of persistence of past trade links differ and the relative importance of various channels is unclear. We show that neither correlated consumption preferences, nor more homogeneous measures explain the persistent results we find. Our results show that even in situations where bilateral trade does not persist, trade composition can be shaped by the former network for centuries. Our results further suggest that trading capital at the level of individual members of a network could contribute to prior findings of trade persistence.

This paper proceeds as follows. In the next section, we provide a brief historic overview. Section 2.3 explains the dataset we assemble for this paper and discusses choices we make in its creation. In Section 2.4, we develop our research hypotheses. Our empirical strategy and aggregate bilateral results are laid out in Section 2.5, followed by our key results in Section 2.6. In Section 2.7, we probe potential mechanisms, and Section 2.8 concludes.

2.2 Historical overview

2.2.1 The Hanseatic League

The Hanseatic league was initiated in the late 12th century in what is now Northern Germany and grew to incorporate the merchants of dozens of cities on the Baltic and North Sea.[8] The formal birth of the League is attributed to the 1161 Gotland Community of German merchants, bringing peace and fostering trade between German merchants and the citizens of the Baltic island of Gotland. Trade initially largely took place among cities on the Baltic Sea, with the cities of Saxony and Westphalia later also represented in the Gotland Community.

This geographic organisation and the role of the League's politicial leaders were continued when the association grew into the Hanseatic League, incorporating more towns and trading increasingly also with the North Sea.[9] The League was primarily a commercial confederation, but it also engaged in common military operations. At its height, the Hansa was a powerful independent political and economic entity. Hanseatic networks played an important role in facilitating exchange by imposing a set of rules and guaranteeing their traders' trading privileges abroad. The traders that were members of the League were called Hanseats. Membership in the League allowed traders to overcome the risk of expropriation and arbitrary taxation. The Hansa built a large and successful transnational trading network. Of great importance were the Hansa's trading posts abroad, called *Kontore* or *Faktoreien*. These hubs sometimes encompassed a whole neighbourhood within a foreign city and had their own jurisdiction. Establishing such trade representations abroad would have been prohibitively costly for small medieval trading firms. While harbour infrastructure such as docks and warehouses lowered exchange frictions unilaterally, legal certainty, similar units of measurement and currencies, and connections among traders lowered trading costs among Hanseats bilaterally. The list of member cities of the Hanseatic League changed over time. Figure 2.2.1 gives an overview of the main Hansa cities based on a list compiled by Dollinger (1964). Table B.B.1 lists both Hansa and Kontor cities.[10]

Dollinger (1964) argues for two dimensions to Hanseatic membership. The first, particularly important until the 14th century during the "Hansa of merchants", were the privileges granted to Hanseatic traders in foreign ports. Treaties with foreign rulers permitted trade to Hanseats, assured them of the legal enforceability of their contracts and protection against arbitrariness

[8] The following discussion is mostly based on Dollinger (1964). Other sources are cited where used.

[9] Unfortunately, we cannot test whether the Gotland Community, as a subset of the later Hanseatic League, shaped trade flows after its dissolution. The Community's members were all located on the Baltic Sea or too far inland, implying that we cannot observe trade among its members, as we only see ships passing through the Sound. See Section 2.3 below.

[10] In both cases, we do not show a list of all Hanseatic cities, but those that meet the criteria for inclusion in our dataset, detailed in Section 2.3.4.

and even encouraged trade by granting tax alleviations. Crucially, trade privileges depended on a trader's citizenship. A non-Hanseat could pay substantially higher taxes than a Hanseat, even when trading the exact same goods on the exact same route. From 1434 onwards, these privileges were confined to traders born in a Hanseatic city. Owing to these privileges, strong links were created between Hanseatic cities and their trading posts abroad, most importantly the Kontore. Figure 2.2.1 displays all cities with a Kontor on a map. The four most important Kontor cities were Novgorod in Russia, Bergen in Norway, London in England and Bruges in what is today Belgium.

The second dimension, rising to importance during the 14th century during the "Hansa of towns", concerned relationships within the network itself. City constitutions, a similar language, privileges within the network and common military and political objectives led to the creation of strong bilateral links between Hansa cities. Gaimster (2005) notes the importance of "social and genealogical links which developed between trading partners and towns and families the length and breadth of northern Europe"(p. 412). Hanseatic trade was usually carried out by self-employed traders or small firms. These firms were too small to establish trading outposts in many cities and therefore relied on networks. Partners could, for example, employ each other as commercial agents. This was often done without written contracts. Instead, the Hanseatic network relied on other coordination mechanisms such as a common language and cultural identity, habits, reciprocity, trust, and reputation. (Ewert & Selzer (2015)) Kinship also helped to make networks more dense: "[I]t was a common strategy of merchants to make occasional partners friends and friends relatives."(Ewert & Selzer (2015), p. 184f.) Migration and return migration further supported this process (Gaimster (2005)). In line with this, Wurpts (2020) uses network analysis to analyze trading partnerships among merchants in Lübeck and finds a strong role of kinship.

The result of these processes was that Hanseatic cities were tied together not only through commercial links but developed a shared culture and economic practices within the network. As part of these, Gaimster (2005) notes the use of lower German dialect, similar urban outlines and shared styles in architecture and design, all alluding "to the shared religious and social values of the urban bourgeois elite of the region."(p.413) In a similar vein, Wurpts et al. (2018) argue that the Hanseatic network also facilitated the spread of the Protestant revolution among its members. They note that "structures of social relations and shared cultural identity enhanced solidarity among Hansards"(p.239). Through the German colonialisation of the Eastern Baltics, the Hansa expanded into non-German speaking territories, leading notably to the foundation of

Riga in Latvia by citizens of Bremen. Hanseats therefore were not confined to being German, yet shared the aforementioned cultural links.

Initially, Hanseatic trade was carried out by men that were simultaneously merchants, ship owners, and ship captains. Over time, however, these roles differentiated. Merchants continued to own ships, but typically in the form of partnerships, diversifying their risk across shares of several ships. At the same time, captains went from being one of these co-owning merchants to professional, hired employees that were contracted by the merchants. Merchants did not anymore accompany their goods, but instead directed their trade operations from home. This differentiation of roles was marked by conflict, as the incentives of the two groups did not coincide anymore. Merchants were worried about their valuable cargo and distrusted the captains and their crews. To counterbalance the difficulty of supervising captains, they had to be married men with families and were liable with their possessions on land. This made absconding and migrating less attractive. In addition, maritime law became increasingly codified. (Dollinger (1964), Jahnke (2017))

In the following centuries, several major trends led to the demise of the Hanseatic League. Competition from Danish, English and especially Dutch merchants meant that Hanseatic traders no longer enjoyed their monopoly power over some of the most lucrative trade routes. The increased competition also laid bare internal conflict between the cities, for example in the question of whether to exclude Dutch merchants from the Baltic or not. Similarly, when England revoked Hanseatic privileges in the mid-16th century, Hamburg betrayed the League by granting English merchants special privileges. Based on voluntary cooperation rather than centralized power, the Hanseatic League also could not compete with the fiscal and military strength of the early modern states in Europe and struggled to form a common policy during the Thirty Years War. More generally, it is well documented that the economic centre of gravity moved towards Western Europe (Acemoglu et al. (2005)). Lacking the support of a powerful state, Hanseatic merchants played only a very limited role in the increased trade with the colonies that caused this westward movement.

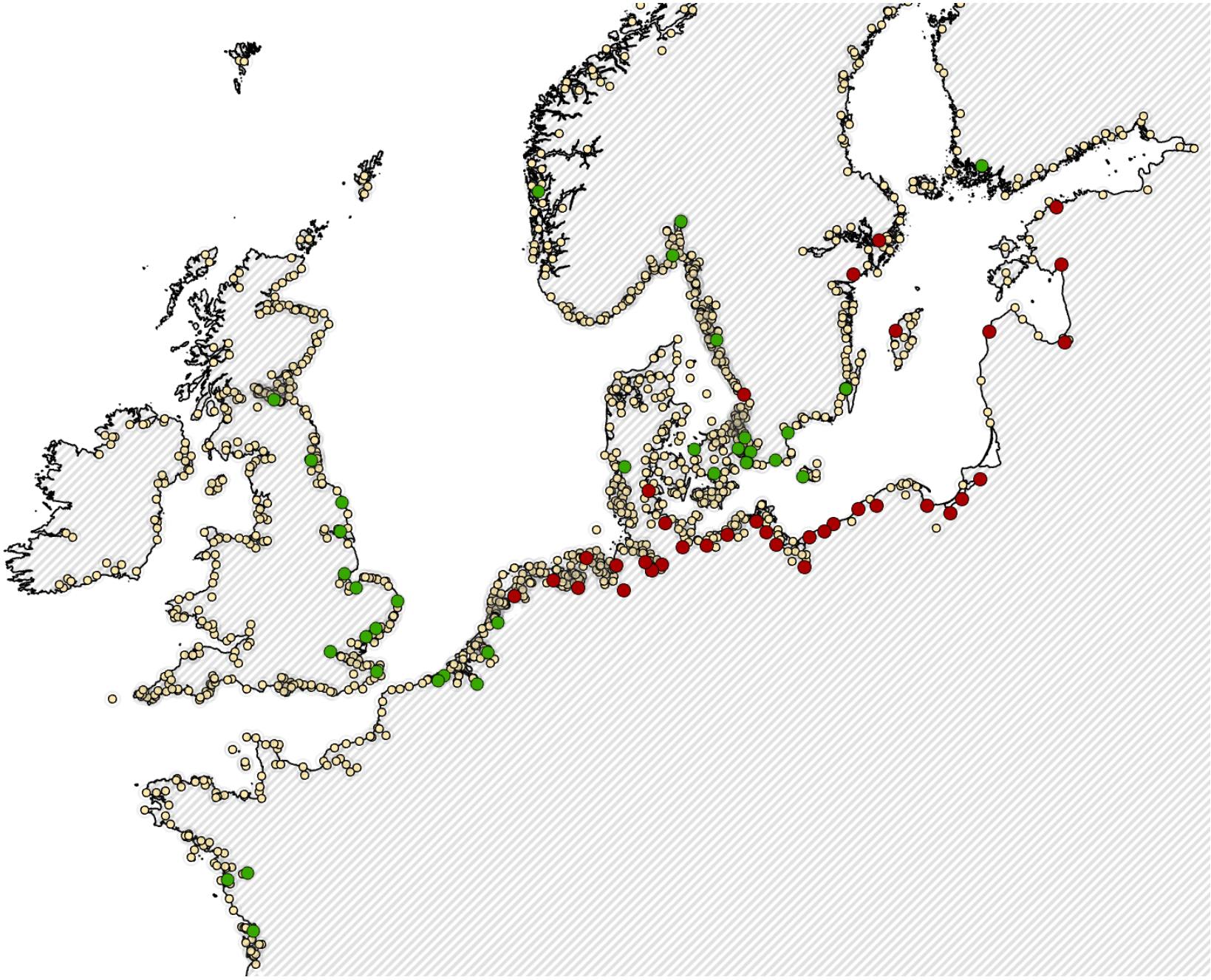


Figure 2.2.1: Red circles show Hansa cities according to Dollinger (1964), green circles show Kontore according to Hammel-Kiesow (2000); all other cities and towns in yellow.

2.2.2 The Danish Sound Toll

To measure trade flows between cities in the Baltic region, we rely on Danish toll data. Introduced by the Danish King in 1429, the Sound Toll was levied on all ships entering or leaving the Baltic Sea at Helsingor (henceforth "the Sound"). At the time, both sides of the Sound were Danish, and the narrow strait between Denmark and what is now Sweden was protected by a fortress that would attack vessels unwilling to pay the duties. While Danish (and initially Swedish) ships paid lower fees, they were still recorded. Passing the Sound unnoticed was essentially impossible due to the narrowness of the strait.

The toll had flat fees and taxes disproportional to the values declared by captains ((Degn 2017, Chapter 5)). The data are extensive, covering the "homeport", the name of the captain, goods aboard, and tolls levied, as well as the port of departure of the ship (consistently since 1560) and its destination (consistently since 1668). There is a debate over the precise meaning of the recorded home port [11]: It either denotes the home port of a ship or the place of residence of its captain and was used to assign toll rates to ships or captains from different countries and regions. (Degn (2017)) For our analysis, it is not important which of the two interpretations is correct, as long as the variable indicates close ties to the city, which we find likely. In the text, however, we follow the interpretation of Veluwenkamp et al. (2021) that the shipmaster generally lived in the town recorded and therefore refer to it as "captain's residence" or "home town". We think of this variable as indicative of the network membership of the captain.

The aforementioned privileges for Hanseatic traders no longer extended to the treatment of their vessels in the period we study.[12] Toll inspections were abolished for Dutch and Swedish ships and similar treaties followed with all major trading partners including the Hanseatic cities, all stating the same toll conditions. From this time onwards, no further major changes in the toll system were introduced until its abolition in 1857. The toll's abolition was a result of diplomatic pressure from Britain and the United States. As our data set begins in 1668, the data are collected at a time when Hanseatic traders enjoyed no privileges at the Sound compared to traders from other nations; the toll conditions did not change in any comprehensive way throughout the period study (1668 to 1857).

[11] The Dutch team of researchers that digitized the Sound Toll data refers to the variable as "schipper plaatsnaam", i.e. "skipper's place name". Boon (1993) briefly summarizes the evolution of the debate since the 1940s.

[12] In earlier centuries, ((Degn 2017, Chapter 5)) notes a hierarchical system in the collection of duties in which ships from different origins paid different duties. Importantly, this differentiated treatment ended in the 1640s and 1650s, after Denmark had signed a treaty with the United Provinces in 1645 reducing the toll to the level at which it was in the 16th century for Dutch traders.



Figure 2.3.1: This map shows the Øresund, where the Sound tolls were collected and all ships recorded, and the Great and Little Belts, the main alternatives for entering or leaving the Baltic Sea.

2.3 Data

2.3.1 Overview

For the digitised trade data, we rely on the Soundtoll project[13], while additionally drawing on (Degn 2017, Chapter 5). In identifying cities in the trade data, we are indebted to the Soundtoll project, which has linked historic port names to unique identifiers. We convert 90,737 original city names to 3,085 unique place identifiers. Other sources are indicated.

An important limitation of these data is that only ships passing the Øresund are recorded. This strait between Denmark and Sweden is displayed in Figure 2.3.1. Trade between Danzig and Stockholm or London and Amsterdam will not feature in this dataset, yet trade between Danzig and London will be observed. We only include trade flows that should be observed, such that not observing a trade flow most likely implies a true zero.

As we explain in more detail in Appendix B.A, it is highly unlikely that any trade between Danzig and London occurred other than through the Sound. Alternative routes were unsuitable and also faced taxes.[14] Tax records confirm that alternative routes played a very minor role (Bergsøe (1853), Gøbel (2010), Degn (2017)).

In the vast majority of cases, the Sound Toll records one port of departure, and one port

[13]http://soundtoll.nl/ We work with a version of the full dataset which we downloaded on November 23rd 2019.

[14]One exception may be the Stecknitz Canal, constructed in the late 14th century to connect the Baltic Sea in Lübeck to the Elbe near Hamburg. The Canal was relatively flat, vessels had to have a draft of no more than 0.43m. The prime commodity shipped was salt, and since the 15th century, commerce was controlled by merchants from Lübeck and their hired captains, who formed their own guild. (Scheftel (1999)). This renders our estimates a lower bound of the Hanseatic network effect as some of the network trade would have passed through this canal.

of destination. We of course cannot rule out multi-strop trips, for example from Bremen to Hamburg, thence across the Strait to Danzig and onwards to Riga. In these cases, it is likely that the departure and origin recorded refer to the last stop before and first stop after the Strait. The presence of trip-chains would thus lead us to observe cities closer to the Sound at a relatively high frequency, but this is not the case. Only very rarely (412 out of 2.2 million voyages) are several cities mentioned on either side and we drop these observations. Similarly, we drop observations for which several captains's residences are recorded, and whenever any city is denoted by "Unknown". Appendix B.A holds the details of this discussion. In addition, the rich sets of fixed effects we use in our main estimations, including city-year and bilateral fixed effects, would absorb much of this type of selection bias.

2.3.2 Measuring trade flows

We measure trade flows between two cities in a given time period by the number of ships that sail from city A to B. While more granular data were recorded, historical units suffer from a lack of comparability, overlapping names in the data, and poor documentation.[15] Aggregation over all kinds of goods and several centuries and languages would lead to high degrees of noise, making the number of ships our preferred measure of bilateral trade.

To alleviate concerns that ship sizes vary by the origin of the captain, we include (captain's residence $\times$ time) fixed effects throughout in our analysis below. Furthermore, we use tonnage information, which is recorded for about 17% of ships, and ask whether ships with a Hanseatic captain differed in size. Appendix Figure B.B.1 finds no evidence for this. Counting ships has the further advantage of being robust against fraudulent or missing reporting of cargo. (Degn 2017, Chapter 5) describes the number of ships passing the Sound as a reliable variable. The Sound is narrow and passing through it unobserved is almost impossible.

In robustness checks, we use tax (or toll) value paid by the ship as the dependent variable. Value is recorded for about 70 per cent of voyages and 80 per cent of entries included in our panel. We use information on currency composition provided by the Soundtoll team and in the rare case of foreign gold coins convert these to Danish gold coins based on their respective gold content. The amount of taxes paid is a proxy for the value of goods shipped. Gøbel (2010) elaborates on the proportionality of these duties, which Waldinger (2022) summarises as a duty of 1-2% of cargo value.

[15] Examples for the problems in harmonizing historical units are that a ton from Danzig is not the same as a ton from Königsberg, a Faad is 930 liters or 950 kg, depending on the good, and many units are abbreviated "D." or "F." leaving ambiguity whether this refers to the currencies of Daler and Faad or other units of measurement.

Throughout our analysis, we code trade flows in a directed way, such that flows from A to B and flows from B to A are distinct observations. In our captain's residence-level analysis, we further distinguish trade flows from A to B by the home town of the ship's captain. This is to say that trade flows from A to B on ships with captains from town C are a separate observation from trade flows from A to B on ships with captains from town D. We control for own trade, that is for trade from A to B on ships with captains from either A or B, to avoid picking up mechanical own trade effects, prompting captains from A to sail more from or to A. Instead, we seek to identify the network effect: Did Hanseatic cities A and B predominantly have their trade facilitated by captains from another Hanseatic city?

2.3.3 Hanseatic cities and the Hanseatic network

We call a route "Hanseatic" when it connects Hanseatic cities, Hansa and Kontor cities, or Kontor cities. This captures the nature of these bilateral connections occurring within the group of Hansa-Kontor cities, displayed in Figure 2.2.1. When adding the membership of captains in the Hanseatic League, proxied for by either their home town or their last name, we will speak of the Hanseatic network. This therefore encompasses trade on Hanseatic routes carried out by Hanseatic captains. For captains, we will not treat Hansa and Kontor citizens symmetrically, as citizens of Kontor cities were considered foreigners and not eligible for Hanseatic privileges. We do not observe captains' entire network but instead proxy for their membership in the Hanseatic network by looking at their home town or their last name.

Cities' membership in the Hanseatic league is determined based on Dollinger (1964). In his accounts, he explicitly includes cities as Hansa members under the condition that their citizens benefited from Hanseatic privileges abroad. The classification of Kontor cities is based on Hammel-Kiesow (2000).[16] A list of Hansa and Kontor cities is provided in Table B.B.1.

As discussed above, it is still under debate whether the home recorded in the Soundtoll data refers to the home port of the ship or the home town of the captain. For our analysis, it is not important which of the two interpretations is correct, as long as the variable indicates close ties to the city, which we find likely. As captains were bound to their home ports in numerous ways, we consider it plausible to interpret a Hanseatic residence as an indicator of membership in the Hanseatic network. As an alternative proxy for being a Hanseatic captain, we use captains' last names, which the Soundtoll Data list throughout. To assess whether captains were Hanseatic, we gather data on the last names of eleven Hansa cities' mayors up to 1700 and add all recorded last

[16]We use a different source here, as Dollinger (1964) does not provide a list of Kontor cities.

names for attendees of the League's general assemblies.[17] A last name is deemed Hanseatic when exactly coinciding with a mayor's or general assembly attendee's last name, which is, however, only a lower bound on classifying Hanseatic last names. In total, we assemble 797 unique last names from mayors' lists and 104 unique last names from Hansetag attendees. Thus, 901 out of 23,108 last names in our data are coded as Hanseatic, covering 3.5% of trade flows. This compares to 21% of trade flows classified as Hanseatic when basing the classification on the captain's home town. Despite overlapping with German last names, we argue in Section 2.2 that being a Hanseat does not coincide with being German. In line with this, many of the Hanseatic last names exist in Dutch and the Scandinavian languages. Rather than excluding individual last names on these grounds, we prefer to rely on our data-driven approach based on the last names of Hanseatic mayors and diplomats.

2.3.4 Dataset

The time frame of our analysis starts in 1668, from which on destination cities were systematically recorded. This is one year before the Hanseatic League gathered for its last diet in 1669. Our trade data end in 1857, when the Soundtoll was formally abolished. Our model below will be fully time-varying, implying that parameter estimates for individual time periods are robust to dropping earlier or later years.

We restrict our sample geographically to coastal areas in Northern and Western Europe. Appendix B.A describes the geographical selection criteria and their implications for our dataset. Virtually all trade flows are between cities on the coast in the Northern half of Europe. These restrictions leave us with 1,425 cities in 16 modern day countries. Among these 1,425 cities, we observe 17,648 bilateral trade relationships over 190 years, with about 1.5 million passages. Accounting for the fact that only trade connections which cross the Sound are theoretically observable to the researcher would yield 485,616 theoretically possible city pairs, of which the vast majority is not observed. Zeroes are common in trade flow datasets, amplified in our case by the granularity of the data, which are at the level of towns rather than countries. For our bilateral analysis, we create a balanced sample of all possible bilateral connections. Missing values for a given time period and city pair mean that in this time period, no passage from A to B across the Øresund was recorded. Hence, we set these missing values to 0. This is to say that we impose stability for a maximum set of cities. Unfortunately, the same approach is not possible

[17] The cities whose mayors' names we record are Lübeck, Hamburg, Bremen, Rostock, Greifswald, Danzig, Stettin, Stralsund, Anklam, Demmin, and Reval (Tallinn). Mayors' names have been collected from Wikipedia lists. Hansetag attendees compiled from Waitz (1870).

for our trilateral dataset that consists of origin, destination, and captain's home town. With 1,425 cities, there are almost 700 million possible permutations (485,616×1,425) per time period. Trying to fill in the entire matrix of all possible trade connections is computationally infeasible. We therefore exclude combinations of origin, destination, and homeport that we never observe over our 190 years of analysis. We then address the potential selection of routes by incorporating origin-destination-time fixed effects. This absorbs never observed origin-destination pairs, as these would show trade flows of 0 for any captain's residence.

Other selection problems are created by the fact that the Sound Toll data only features trade flows which enter or leave the Baltic Sea through the Sound near Helsingor. On short distances around the Sound, we likely miss trade, as land transportation was a viable alternative. As a robustness check, we exclude short-distance connections for which we likely observe only part of trade due to land transport. Some selection issues are also alleviated by the fact that we use a balanced panel of trade connections.[18] Moreover, our main specification includes (origin × destination × time) fixed effects which account for time-varying bilateral selection.

Between the cities in our sample, we compute cost distances for each city pair using a raster approach and the CostDistance tool in ArcGIS, similar to Nunn & Puga (2012) and Bakker et al. (2021). As in the former paper, our pixel resolution is 30 arc-seconds, corresponding to square cells of about 1 km side length. We compute 1,425 × 1,424 distances for our sample of coastal cities. Most least-cost routes are over water, but we occasionally require short journeys over land when cities are located a bit inland. In those cases, we use a relative cost parameter of 10 for land transport over water transport as in Michaels et al. (2012) and Maurer & Rauch (2022). Table 2.8.1 holds key descriptives on the period we cover, the number of entries, ships, and cities, the coverage for tax value and last names, and the number of observations in the trade panels we construct.

2.4 Hypotheses

There is ample empirical evidence that shared institutions, languages and currencies increase trade flows (Jacks et al. (2020)), and the same holds for trust (Guiso et al. (2009)). Previous studies have also pointed out that the "trading capital"generated by such institutions or informal networks depreciates once the underlying institutions and networks disappear. There is little

[18]We further address endogenous zeros by restricting estimation to a sample where observations are far from the truncation point. Specifically, we only include connections which are used at least 60 times throughout the 190 years of our data, so about one in three years. Our analysis on this sub-sample of frequently used connections produces virtually identical results. These results are available upon request.

evidence as to how long this depreciation may take. To answer this question, we study the former network of Hanseatic traders and their intermediaries. Our main hypothesis is that Hanseatic traders had accumulated particularly large stocks of trading capital on Hanseatic routes. Therefore, the depreciation of this trading capital should have taken a longer time compared to other traders on the same routes. This would be reflected empirically in Hanseats trading significantly more within their networks, i.e. on Hanseatic routes and using Hanseatic intermediaries.

We begin by applying the notion of trading capital to the Hanseatic League. Beestermöller & Rauch (2018) classify the "trading capital"created by past interactions into three components: physical capital like infrastructure that reduces trade costs; capital related to personal networks and built-up trust; and other components that lower bilateral trade costs, but are unrelated to personal networks or physical capital. While we do not study road and railway construction, the accumulation of physical capital matters in our setting. Close trade links resulted in the construction of representations, warehouses, docks, and cranes and this accumulated physical capital may have continued to reduce bilateral trade costs after the dissolution of the League. General harbour investments, however, are to be viewed as a public good, reducing the destination cost component for all origins. The buildings representing the Hanseatic League reduced trade costs only for traders from Hansa cities. The contribution of this physical capital should have ceased, however, when there was formally no League to be represented any more. We conclude that physical capital will have mattered for the accumulation of trading capital, but that it was insufficiently discriminatory to generate the network effects we focus on.

The second category of trading capital contains personal networks and built-up trust. Both are likely to have outlived the Hanseatic League. As argued in Section 2.2, Hanseatic trade relied heavily on networks, reciprocity and trust, and we expect this category of trading capital to be particularly important in our setting, especially given the distrust between merchants and captains. The last category encompasses cultural proximity, shared languages and trading habits. From our point of view, the last two categories can hardly be distinguished precisely.

In the context of the Hanseatic League, trading capital consisted, among others, of a more symmetric distribution of information, well-aligned legal systems, similar units of measurement, a shared language and a close-knit network of traders and captains. Hanseatic networks were marked by kinship and traders pursued family connections as an informal means of accumulating trading capital. Wubs-Mrozewicz (2017) stresses the importance of the Hanseatic league as an institution of conflict management with formalized mediation mechanisms for conflict among

individuals and conflict among cities. Such mechanisms and other legal principles, such as the symmetry of Hanseatic privileges abroad for all members, outlasted the League as concepts in law and are likely to have lowered trade costs for involved cities even after the dissolution of the league. Therefore, trading capital was built up through these legal institutions. Privileges abroad created trading capital in the first era of the League's existence (Dollinger (1964)). Lower bilateral trade frictions and political and cultural similarities led to the creation of trading capital among Hanseatic cities in the latter phase of the League's existence. Both privileges abroad, as these were historically linked to holding a Hansa citizenship, and cultural similarities between traders and captains suggest an important role for individuals. This motivates our hypothesis that trading capital had been accumulated at the level of individual traders and captains and their respective families.

Our hypothesis is that the accumulation of trading capital varied for historical reasons. We propose a mechanism in which social connectedness in networks, similar legal and economic institutions, and trust through repeated interactions accumulated trading capital between Hanseatic traders and captains. The fiscal and legal privileges conveyed by the Hanseatic League upon its members were linked to a trader's citizenship: only Hanseats benefited from lower tax rates, legal certainty, and military protection. Therefore, our hypothesis is that this gave Hanseatic merchants and captains particularly large incentives to invest into the accumulation of trading capital on Hanseatic routes. The returns on investment on Hanseatic routes were lower for non-Hanseatic traders than for Hanseats, given that non-Hanseats did not receive the financial and legal benefits Hanseats received when trading on Hanseatic routes. On the other hand, distrust between Hanseatic merchants and Hanseatic captains was lower due to cultural similarities (and maybe even extended kinship networks). Thus, we expect particularly large stocks of trading capital within the Hanseatic network.

Given this discussion, we expect a positive, but declining effect of common membership in the Hansa on trade. Excess trade flows due to trading capital would thus go to zero over time. In the case of Head et al. (2010), this process took only a couple of decades. However, our trade data begin only in 1668, one year before the formal dissolution of the Hanseatic League and already after a process of gradual decline, with many cities having dropped out informally. It may therefore be the case that the city-level trading capital had already eroded by the time we start observing trade flows. However, based on our above hypothesis, we expect a longer lasting effect of a route being Hanseatic conditional on captains being Hanseats. We thus expect trade flows in excess of gravity forces for a long time period if the origin and destination cities and the

facilitators of trade were part of the former Hanseatic network.

2.5 Descriptives and bilateral trade data

We start our analysis by giving some descriptives. In Figure 2.8.1a below, we plot the share of trips through the Sound that originated in a Hanseatic city (left axis) and the number of trips registered at the Sound Toll station (right axis). As can be seen, the Hansa had been losing ground already before its dissolution in 1669, and this trend continued afterwards. This decline could in theory be explained by a general relative decline of the importance of Hansa cities. However, in Appendix Tables B.B.2 and B.B.3 we show that Hansa cities were neither larger nor smaller than other cities in our dataset and seem to have remained wealthier, as expressed by higher levels of imports and higher import shares of luxury goods. Figure 2.8.1b shows that Hansa cities became relatively less active in the facilitation of trade, as indicated by the declining share of captains from Hanseatic towns. Figure 2.8.1c shows that following its dissolution, trade within the League as a share out of all trade declined from 60% in 1700 to 20% in 1800. Thus, both the overall volume of trade on Hanseatic routes and the share of trade carried out by captains from Hansa towns port indicate the reduced importance of the League. Nonetheless, former Hansa cities remained key players throughout most of the 18th and 19th century, being the origin of between 20 and 40% of all the trips through the Sound, as seen in Figure 2.8.1a. This is particularly true for trade on Hanseatic routes: Figure 2.8.1d shows that on average 40% of trade flows on Hanseatic routes were carried out by Hanseatic captains. This share does not display the declining trend seen in Figure 2.8.1b for trade overall. Thus, despite the League's dissolution in 1669 and its evidently reduced importance, Hanseatic trade continued.[19]

We next examine whether the Hanseatic League persisted in city-to-city trade flows. For this, we draw on a bilateral dataset that consists of 485,616 directed trade flows, each over 190 years. These are aggregated to 20 decades.[20] Therefore, we count a total of 9,712,320 observations. In our balanced panel, 99 per cent are zeros. This compares to roughly 80 per cent zeros in Jacks et al. (2020) who study the Empire's trade enhancing effect in British bilateral export data at the country and commodity level. As our data are city-level and we have far more cities than Jacks et al. (2020) has commodities, a higher share of zeros is to be expected. Descriptives for

[19]These graphs differ in the periods they cover due to data limitations. Since we only require the origin in Figure 2.8.1a, we can plot this series from 1537 onwards. Between 1497 and 1537, only a handful of ships are recorded. In our main analysis, we require data on both origin and destination, which limits our analysis to the period between 1668 and 1857. As evident in Figures 2.8.1c and 2.8.1d, few destinations are recorded before 1668. Home towns are systematically recorded throughout, as seen in Figure 2.8.1b.

[20]The first and the last one do not count ten years. We aggregate to decades rather than using the annual data due to computational limitations encountered when adding the dimension of a captain's home town below.

this bilateral trade dataset can be found in Table 2.8.1. Our empirical model for this dataset follows standard gravity models:

$$T_{ijt} = exp(\alpha + \beta_t ODHansa_{ij} + \lambda_t dis_{ij} + \eta_{it} + \chi_{jt}) \times \epsilon_{ijt}, \tag{2.5.1}$$

where T_{ijt} denotes the number of ships sailing from city i to city j at time t, our measure of trade between cities. $ODHansa$ is a dummy that codes whether the origin and destination city formed a Hanseatic route, meaning they were either a Hanseatic or a Kontor city. The coefficient for this dummy indicates by how much aggregate city-to-city trade flows are beyond or below the prediction of the other gravity factors. The dummies are time-invariant, so we interact them with a complete set of decade dummies to see how the effect of Hanseatic trade linkages evolves over time. As a control variable, we include the log distance over sea between the two harbours, dis_{ij}, also interacted with decade dummies. η_{it} and χ_{jt} are origin x decade and destination x decade fixed effects. As shown by Fally (2015), estimating a gravity equation with PPML and including importer and exporter fixed effects solves the problem of multilateral resistance terms pointed out by Anderson & van Wincoop (2003). ϵ_{ijt} is the error term. We estimate this equation on the bilateral directed dataset and use PPML[21] (Santos-Silva & Tenreyro (2006)), clustering standard errors on origin-destination pairs throughout to allow for serial correlation at the bilateral level.

Our main result from this estimation are 20 decadal parameters for the Hanseatic dummy from equation 2.5.1. Point estimates and their 95% confidence intervals are shown in Figure 2.8.2. Appendix Table B.B.4 presents all regression details. Persistence in aggregate bilateral trade flows would manifest itself in statistically significant and positive coefficients in the decades following the end of the Hanseatic league, in line with Head et al. (2010) or Beestermöller & Rauch (2018). As can be seen, we do not find such persistence. The coefficients are usually statistically insignificant and close to zero.[22] The only exception to this pattern are the first decades of the 18th century, when the coefficients even turn negative. This is likely due to the severe trade disruptions in the wake of the Great Northern War that disproportionately hit trade among former Hansa cities.[23] The estimated coefficients turn negative at the beginning of this

[21]The specific Stata package used throughout is PPML HDFE by Correia et al. (2019).

[22]Through a simulation exercise, we rule out that this is driven by sample selection. The Soundtoll Data only include trade flows that pass through the Sound. Therefore, we run 100 simulations of all trade flows and impose a positive effect for $ODHansa_{ij}$ by setting $\beta = 1$. We then run the PPML estimation of equation 2.5.1 on the entire simulated sample and on the sub-sample that would be observed, namely all flows that pass through the Sound, for the 100 simulated trade flows. The histogram of estimates on $ODHansa_{ij}$ is presented in Appendix Figure B.B.3. The mean value of β estimated in the simulated sample is 0.999986 with a standard deviation of 0.002334, which compares to a mean estimate of 1.0023 with a standard deviation of 0.00315 in the observed sub-sample. Each estimate is within the 95% confidence interval of the other estimate. In both cases, the simulated β coefficient on $ODHansa_{ij}$ is in the range of supported estimates and no bias is detected as arising from this sample selection.

[23]While few Hansa cities were subject to direct war action, the war was associated with a severe

conflict in 1700 and reach their lowest point in the 1720s, when the war ended. The recovery to pre-war levels of insignificance takes about three decades. This is longer than the 8 years found by the study on war and trade by Glick & Taylor (2010), though it should be noted that their study only extends back to 1870, when greater economic integration may have already reduced the costs of war. Further, the Great Northern War's effects were aggravated by a severe plague outbreak (Marczinek (2024)). Apart from these decades of war and recovery, trade volumes between Hanseatic cities are similar to those of other routes across the Sound.[24] In the appendix we show robustness estimations for this result.[25]

The aggregate bilateral specification allows to check for persistence in aggregate trade patterns similar to Head et al. (2010) or Beestermöller & Rauch (2018). However, it does not allow us to dissect the identity of the traders and their intermediaries. We therefore next turn to our main analysis that incorporates the captain's home towns.

2.6 Results

2.6.1 Main Results

Our results so far disregard the network of traders and captains. As the historical discussion has shown, traders' network membership was important in medieval Baltic trade, and distrust between merchants and captains was rife. While we cannot ascertain a trader's entire network, we argue that a captain's residence is a good proxy for him belonging to the Hanseatic network. Our alternative proxy is the captain's last name. We expect captains from Hanseatic cities to have accumulated more trading capital than others when trading on Hanseatic routes.

outbreak of the plague, which afflicted almost half of all Hansa cities, and in particular important Hansa cities such as Stettin, Königsberg, Elbing, Hamburg, Stralsund, Greifswald, Pernau, Kiel, Reval, Visby, Riga, Malmö, and Danzig. The Southern and Eastern Baltic were at the centre of the plague outbreak and most Hansa towns were located there. (Kroll (2006)) Exports even of some of the larger Hansa cities drop to close to zero during this period. However, as the entire Baltic region was involved in the war and the Hanseatic League was not an active political entity any more, there is no evidence that Hanseatic cities were targeted by either the war or the plague.

[24]The Napoleonic Blockade coincides with the period and region we study, but does not influence results strongly as it did not affect Hanseatic cities proportionally.

[25]For the bilateral dataset, we show that also the value of trade does not show any persistent effect of the Hanse (Figure B.B.4). We establish robustness to PPML by showing results from using OLS and trade in logs (Figure B.B.5). Furthermore, we show that time or intensity variation in membership of the League does not matter. Even among the cities attending the League's last general assembly in 1669 (Bremen, Hamburg, Lübeck, Rostock, Danzig) no trade persistence can be found (Figure B.B.6a), with results virtually identical to those in Figure 2.8.2. Similarly, we use data on Hansetag attendance between 1492 and 1516 from Poeck (2010) (also used by Wurpts et al. (2018)).[26] A total of 49 cities in our sample actively attended the League's general assemblies in these years, which we interpret as the most active members. Results are shown in Figure B.B.6b and again virtually identical to those in Figure 2.8.2. Finally, we present results from an aggregated specification without any time dimension, on which we can also apply the bias-correction for PPML suggested by Weidner & Zylkin (2021) (Table B.B.5). The aggregated specification also allows us to more easily compare results across different specifications and robustness checks.

We therefore add captains' home towns to our analysis, which we argued above is a proxy for a captain belonging to the Hanseatic network. In the resulting dataset, an observation is the number of ships sailing from city A to city B owned by captains in city C in a decade d.

Our key variable of interest is $AllHansa_{ijn}$, a dummy that is one when the origin and destination were either Hansa or Kontor towns, and when the home town was a Hansa town. For origin and destination, we continue to treat Hansa and Kontor towns symmetrically, as we are interested in trade flows on Hanseatic routes specifically. Our discussion both of historical trading privileges in Section 2.2 and the notion of trading capital at the level of traders and captains we propose suggests a stricter view on the captain's home town, however. We do not treat Hanseatic and Kontor home towns symmetrically, but insist on a Hanseatic one. Ships with captains from Kontor towns did not belong to the Hanseatic network and have to be regarded as distinct. The triple Hansa dummy thus captures trade within the Hansa network: among Hansa cities, among Kontor cities, and between Hansa and Kontor cities, and carried out by ships with a Hanseatic captain.

To focus on the network effect, we control for a same harbour dummy, $Harbour_{ijn}$, that takes on the value of 1 whenever a captain's home town coincides with the origin or the destination of the journey. As captains are generally more likely to sail from and to their home town, this is an important control variable to rule out mechanical effects. The dependent variable is now T_{ijnt}, denoting the number of ships sailing from city i to city j with captain's home town n at time t. In our main specification, we control for time-varying route characteristics by introducing origin x destination x time fixed effects. This is very important, as it addresses the selection of bilateral routes into our dataset, as well as worries that the rise of the Atlantic trade and the general westward shift of trade drive our estimates on the $AllHansa_{ijn}$ dummy. Furthermore, this controls for possibly correlated consumption preferences among residents of former Hansa cities, as we compare Hansa ships to non-Hansa ships trading between the same towns. To further absorb multilateral resistance terms (Baldwin & Taglioni (2007)), we also include home town x time fixed effects:

$$T_{ijnt} = exp(\alpha + \beta_t AllHansa_{ijn} + \delta_t Harbour_{ijn} + \chi_{ijt} + \xi_{nt}) \times \epsilon_{ijnt}, \tag{2.6.1}$$

where T_{ijnt} denotes the number of ships sailing from city i to city j with home town n at time t, $Harbour_{ijn}$ is a dummy equal to 1 whenever a captain's home town coincides with a journey's origin or destination. ϵ_{ijnt} is the error term. The coefficients β_t on $AllHansa_{ijn}$, a dummy that is one when all of origin, destination and home town are part of the Hanseatic network, are what

we are most interested in, as it measures by how much trade on Hanseatic routes overshoots gravity predictions when executed by Hanseatic captains. $AllHansa_{ijn}$ and $Harbour_{ijn}$ are interacted with decade dummies. χ_{ijt} and ξ_{nt} are origin x destination x decade and home town x decade fixed effects. Note that χ_{ijt} absorbs the bilateral Hansa dummy, $\gamma_t ODHansa_{ij}$, and distance over sea, $\lambda_t dis_{ij}$. Generally, the time-varying bilateral fixed effect absorbs all bilateral variation, in particular distance between origin and destination, commonality variables such as language similarity, the network structure of bilateral trade, and changing bilateral trade patterns due to the steam ship (see for example Pascali (2017)). The origin x destination x time fixed effect also nests origin x time and destination x time fixed effects, accounting for the fact that Figure 2.8.1a below shows that Hansa towns' exports grow more slowly. As χ_{ijt} effectively controls for total trade on a route ij in time t, β_t captures the relative importance of Hansa cities in bilateral city-level trade. The home town x time fixed effect further addresses the issue of a declining Hanseatic role in shipping overall, as documented in Figure 2.8.1b below. While χ_{ijt} absorbs the distance between origin and destination, additionally controlling for decade-interacted distances between home town and origin and between home town and destination produces virtually identical results.[27] These controls capture some of the multilateral logic of trade and the estimated distances are significant and negative throughout. Throughout, standard errors are clustered on origin-destination pairs. This allows errors for different home towns and different years, but on the same bilateral route, to be arbitrarily correlated. Results clustering standard errors on the origin-destination-home town level are virtually identical and available upon request.

Some examples can illustrate the comparison that we are making here: On average, we see 1.3 captains from Hansa city Lübeck doing the trip from the Hansa city Hamburg to the Hansa city Riga per decade. However, on the same route, we only see 0.1 ships per decade with a captain from the non-Hansa city St. Petersburg. As another example involving the Kontor city of London, there are no vessels with captains from Amsterdam that ever sail from Lübeck to London, but we see each decade on average 0.75 ships on that route with a captain from Hamburg. Specification 2.6.1 allows us to examine whether these examples are generalizable, i.e. whether we do see a general pattern that trade between Hanseatic cities is facilitated by Hanseatic intermediators.

Our results are displayed in Figure 2.8.3a and show that this is indeed the case. A pair of cities on a Hanseatic route in combination with a Hanseatic captain implies significant excess trade flows. Appendix Table B.B.6 presents all regression details. Note that the inclusion of

[27] Results available upon request.

$Harbour_{ijn}$ means that these results are not driven by captains from a pair of cities generally sailing more between these two cities, independent of Hansa status. Further, these positive effects are economically large: They average about 0.5 with little time variation, implying excess trade of 64.9%, and amounting to an equivalent tariff reduction of 9.5%.[28] They also persist until well in the 19th century, with a weak downward trend. Thus, even though the Hanseatic League had no lasting effect on aggregate city-to-city trade, it had a strong and long lasting effect on the composition of trade flows operating via individual traders and their facillitators.[29] In terms of magnitudes, the average value of shipments originating in a Hansa or Kontor port on a ship captained by a skipper from a Hanseatic town and going to a non-Hanse and non-Kontor destination is 17.6 Danish dalers per decade. Changing the destination to a Hansa or Kontor one moves this average shipment value up to 75 dalers. In terms of the count of ships, the average is 0.37 per decade for Hansa ships originating in a network port and sailing to a non-network port, and 1.19 for Hansa ships originating in a network port and sailing to another network port. This is even true when conditioning on any trade flow taking place between the cities over a decade, with the averages at 3.8 and 7.6, respectively.

While we argue that a captain's home town is a plausible proxy for being a Hanseat, our results do not depend on this specific approach to identifying captains. In Figure 2.8.3b, we replace the captain's home town dimension by his last name. We compare these last names to a list of Hanseatic mayors' and diplomats' last names in order to judge whether a trade flow was carried out by a Hanseat. We find lower but still significant positive effects of captains being Hanseats. The average estimate in Figure 2.8.3b is half the average estimate in Figure 2.8.3a, which we attribute partially to the much lower share of Hanseatic captains: Only 3.5% of trade flows are identified as Hanseatic based on last names, which compares to 21% when identification as Hanseatic is based on a captain's home town.

2.6.2 Robustness

We show that these results are robust to using the value of the tax paid on the traded goods, rather than the number of ships, as the outcome variable (Appendix Figure B.B.8). This has the

[28] Excess trade is calculated as $exp(\beta) - 1$. Following Anderson & van Wincoop (2004) and Wolf (2009), the tariff equivalent is calculated as $exp(\beta/-\sigma) - 1$. For the elasticity of substitution, we assume $\sigma = 5$, a typical value taken in the literature. Trade costs are trilateral and we control for a bilateral Hansa dummy throughout. Thus, this captures reduced trade costs for Hanseatic captains on Hanseatic routes.

[29] In a placebo exercise in the home town-level analysis, we randomly assign 5% of cities to a placebo Hansa home town, in line with the share of cities in the Hansa-Kontor group according to Table 2.8.1. We repeat this exercise 100 times and run our main model equation 2.6.1, including the same harbour dummies, on each simulated placebo network, with $AllHansa_{ijn}$ defined for each placebo network instead. Histograms for the estimated coefficients on $AllHansa_{ijn}$ are presented in Appendix Figure B.B.7. Collapsed across decades for ease of comparison, the mean estimate, with standard deviation in brackets, is -1.066 (1.041). We conclude that the placebo Hansa membership does not display effects on trade.

advantage of directly measuring the value of trade flows, but was potentially subject to fraudulent underreporting by the captains. As for the aggregate bilateral results, we check robustness to PPML by using trade in logs as an outcome and using OLS (Appendix Figure B.B.9). We further check robustness to the issues of land transport around the Sound and of circumventing the Sound via the Great or Little Belt by excluding short-distance connections (Appendix Figure B.B.10). Moreover, our results are also similar when looking at Hanseatic cities that attended the last general assembly (Appendix Figure B.B.11a) or those that more actively attended general assemblies (Appendix Figure B.B.11b). Similarly, we find no evidence for heterogeneity in the Hanseatic network effect when excluding the five largest Hansa cities (Bremen, Hamburg, Lübeck, Rostock, Danzig - see Appendix Figure B.B.12). This controls for the possibility that the network effect is not driven by unobserved factors that determine city size and trade patterns. We also check that our results are robust to accounting for endogenous zeros by restricting estimation to a sample of frequently used connections that are unlikely to feature zero trade (Appendix Figure B.B.13). We further add as controls time-interacted distances over sea, either between the origin and the home town or the destination and the home town, whichever is shorter (Appendix Figure B.B.14). We also include a dummy for the origin or destination country coinciding with the captain's country of origin to control for a measure of the rise of the nation state. This goes further than our usual control for origin or destination ports coinciding with the captain's home town. As can be seen in Appendix Figure B.B.15, we find very similar results. Thus, we do not simply find that captains sail more from their own home country to another country. As in the bilateral case, we also create an aggregated dataset without time dimension, which allows us to easily compare different specifications and models (Appendix Table B.B.7). Across all these alternative specifications, we find that Hanseatic captains traded significantly more on former Hanseatic routes. This confirms our hypothesis that trading capital on Hanseatic routes was accumulated mostly by Hanseatic traders and captains. Finally, we establish the robustness of our findings to alternative specifications, equivalent to existing gravity models in the multinational production literature (Head & Mayer (2019), Arkolakis et al. (2018)). The estimation equations and results for these specifications are presented and discussed in Appendix B.C.

2.7 Probing Mechanisms

We interpret our results as evidence for the importance of trading capital accumulated by Hanseatic merchants and capaints on Hanseatic routes: Merchants from Hanseatic towns were

more likely to use captains from other Hanseatic towns as facilitators of their trade. As a further corroboration for this, we ask whether Hanseatic captains on former Hanseatic routes were particularly successful in holding on to their market shares. We construct a data set of trade shares in the first years of our sample, 1668 to 1679, for trips where the captain's home town did not coincide with the trip's origin or destination. The median market share of a city A actively trading between cities B and C lies at 12.5% in that period. We construct dummies for having the highest market share and interact this with a dummy for a Hanseatic captain trading on a former Hanseatic route. Similarly, we look at initial market share and interact it with the dummy for a Hanseatic captain on a former Hanseatic route. In Table 2.8.2 we present results of regressing market shares from 1680-1857 on initial market shares, the highest market share dummy, and their interaction with the dummy for Hansa captains on Hansa routes. These results show that, overall, a city that dominated trade on a given route in 1668-1679 is predicted to have a lower market share in the following decades. This decline, however, is relatively weaker for Hanseatic captains on former Hanseatic routes. Within the network, Hanseatic captains may have held on to their competitive advantage over other captains, suggesting a role for the transportation sector as in Brancaccio et al. (2020). As we have excluded observations where a trade route is served by a ship from either the origin or the destination, this result is indicative of a network effect.

An alternative explanation of this finding could be intergenerational transmission of navigational knowledge that disproportionately took place among Hanseats. We find in Table B.B.8 that Hanseatic routes connect more similar cities in terms of ruggedness and temperature. This suggests that they are if anything easier to navigate than other routes, providing evidence against this alternative explanation. Further, the most narrow point of any journey, the Sound itself, has to be crossed on all routes, limiting the variation in required sailing skills. A final argument against this alternative explanation is provided by Figure 2.8.1d: Hanseatic routes were increasingly operated by non-Hanseatic captains, implying that these captains should also accumulate the required sea faring knowledge. After almost two centuries of both Hanseats and non-Hanseats sailing on Hanseatic routes, it appears implausible that only Hanseats would have learned how to sail on these routes.

We now differentiate Hansa-specific trading capital on Hanseatic routes from Hansa-specific trading capital in Hanseatic origins or destinations. The latter encompasses, in particular, physical capital in harbours exclusively used by Hanseatic traders and captains, such as docks, warehouses, and trade representations that would have lowered shipping costs from and to

these harbours but only for Hanseats. To investigate this, we extend model 2.6.1 by including origin-home town-time and alternatively or additionally destination-home town-time fixed effects. For ease of comparison, we collapse the data across decades. These additional fixed effects absorb trading capital specific to Hanseatic captains in individual harbours. Appendix Table B.B.9 holds the results. We find that the estimated coefficients attenuate, but remain significant and positive. The average estimated tariff equivalent of the Hanseatic network effect lies at 9.8%. When including origin x home town or destination x home town fixed effects, the average estimated tariff equivalent drops to 8.8%. When including both origin x home town and destination x home town fixed effect, the tariff equivalent drops to 3%. The finding is still significant. Overall, these results suggest that Hansa-specific trading capital in harbours played a role, but it does not overrule our findings on the network effect.

We next attempt to distinguish different components of trading capital at the level of individual traders and captains. One factor that is stressed by the historical literature (Dollinger (1964)) is that Hanseatic traders used a more homogenised set of units and measures. This would have lowered bilateral trade costs specifically for those Hanseatic traders, in particular when keeping in mind our discussion in Section 2.3 about the variety and idiosyncrasy of historical units of account. We clean these units to the best of our abilities and construct a dataset of 59 historical weight, volume, and count measures. As goods may vary in the variety and dimension of applicable units, we also clean the good information and code 227 different commodities. 90% of observations are neatly mapped onto a cleaned commodity and 97% of the remaining 10% are empty ships, undefined goods, and mercery. Two variables capture the variety of units used: first, the standard deviation of units' shares by origin-destination-home town-decade observation. Whenever only one unit is used, this takes the value of 0. Second, we count how many units are used per observation. In Table B.B.10 we show results from regressing these two measures on a dummy for a Hansa captain on a Hansa route. We find no evidence that Hanseatic captains used a smaller number or more homogenised set of units - if anything, our evidence points to the opposite. This suggests that harmonised units of account did not contribute to the trading capital that we see.

Using the cleaned goods information, we construct similar measures of commodity concentration, i.e. the standard deviation of goods' shares and the number of goods shipped. In Table B.B.11, we find that Hanseatic captains on Hanseatic routes shipped a larger variety of goods. We interpret this as evidence for greater integration in the Hanseatic network and as evidence against the hypothesis of correlated consumption preferences. This result also speaks against a

competing explanation, namely decaying Hanseatic monopoly power in a few goods. We show that, to the contrary, Hanseatic trade was characterized by a larger set of goods.[30] Trading a more varied basket of goods may indicate higher levels of trust between Hanseatic merchants and captains, supported by our finding that this result is not driven by more standardised units.

We have argued that our empirical findings are not explained by navigational knowledge and do not find evidence that this trading capital consisted of more similar units employed by Hanseatic captains on Hanseatic routes. Instead, we show that trading capital was largely specific to Hanseatic captains on Hanseatic routes rather than in individual Hanseatic harbours. Similarly, we show in Appendix Figure B.B.16 that we find no persistence when only the home town and either the origin and destination (and not both) were Hanseatic. We interpret this as the network effect: captains that belonged to the Hanseatic network facilitated significantly more trade on routes connecting this network. Our finding in Figure 2.8.3b, identifying individual captains as Hanseats by their last names, comes closest to testing this hypothesis at the level of individuals. We find that captains with Hanseatic last names traded significantly more on routes within the network. This suggests that trading capital had been accumulated at the level of individual traders and captains, or their families. While we cannot explicitly test for this, this result is consistent with the existence of a Hanseatic identity that persisted in the network.

2.8 Conclusion

In this paper, we employ an extensive city-level trade panel to investigate the long-run behaviour and stability of social and economic networks. We study trade for 190 years after the demise of the Hanseatic League, a confederation of traders in Northern Europe. After the League was dissolved in 1669, we find that former Hanseatic cities did not trade significantly more with each other. All bilateral trading capital that had been created between member cities seems to have depreciated completely by the time the League was formally dissolved.

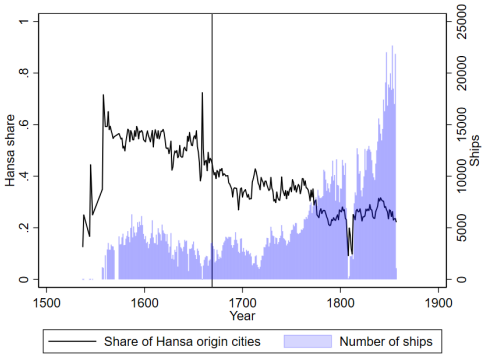
Rather than focusing on Hanseatic cities, we then analyse the network of traders and captains. The quality and high resolution of our data allow us to distinguish captains by their home town and last name. Guided by the historical narrative, which highlights the role of maritime networks and privileges granted to traders from Hanseatic cities, and the importance of trust between traders and captains, we incorporate the captains' network membership into our analysis. Our hypothesis is that Hanseatic captains and traders built particularly large stocks of trading capital

[30] In results available upon request, we further find significant positive estimates in each decade, again speaking against decaying monopoly power driving our result.

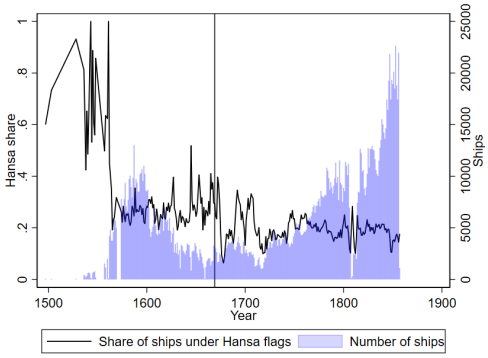
on Hanseatic routes. We determine network membership either by captains' home towns or by comparing captains' last names to a list of Hanseatic mayors' and diplomats' last names. We find that Hanseatic captains continued to trade significantly more on routes within the former network. This result holds more than a century after the end of the League and is established by estimating different gravity models and robust to several measures of trade and sample restrictions. Our finding illustrates the dynamic gains from networks, as they continued to shape trade flows for over a century. Networks have been shown to promote trade and development by lowering frictions. Our results on the long run stability of such networks underscore the scope of these findings. Social and economic networks are hard and costly to build, but once established they can survive for centuries.

Figures and Tables

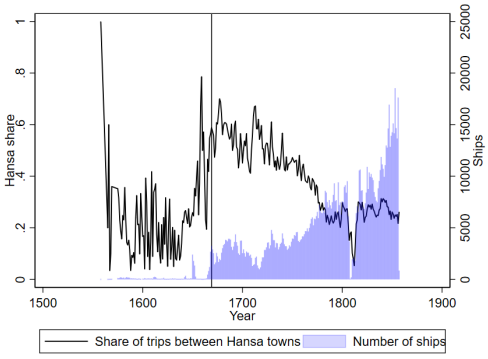
Figure 2.8.1: Hansa trade over time



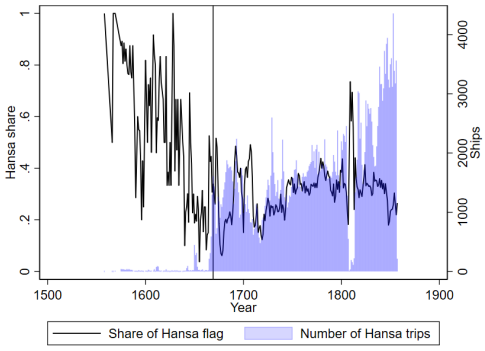
(a) Share of trips originating in a Hanseatic city and # of all ships recorded.



(b) Share of trips on ships with captains from Hanseatic towns and # of all ships.



(c) Share of trips on Hanseatic routes and # of all ships recorded in the dataset.



(d) Share of ships with captains from Hanseatic towns on Hansa routes, # ships on Hansa routes.

Notes: Descriptives on trade composition and Hansa share over time.

Table 2.8.1: Key descriptives

Characteristic	Value	Comment
Coverage, origin information	1560-1857	first in 1497, but unsystematically
Coverage, destination information	1668-1857	first in 1536, but only for fraction of passages
Coverage, sample	1668-1857	we require destination information
# of entries	3,503,612	on average, three goods per ship entered
# of ships	1,205,686	
# of value entries	2,678,236	about 80% of entries
# of ships with value entries	814,650	about 70% of ships
# of cities in the data	3,085	90,737 spellings mapped onto 3,085 unique cities
# of cities in the sample	1,425	See 2.3.4 for sample restrictions
# of home towns	1,059	
# of Hansa and Kontor cities	71	Dollinger (1964), Hammel-Kiesow (2000)
% of Hansa and Kontor cities	5	
% of Hansa cities in entire region's urban population, 1700 (1800)	14 (11)	Population data for 273 out of 1425 cities from Bairoch et al. (1988)
# of Hansa home towns	37	Dollinger (1964)
% of trips on ships with Hansa home towns	21	
% of trips outside Hansa network by ships with Hansa home towns	16	
% of trips within Hansa network by ships with Hansa home towns	33.5	
# of last names	23,108	
# of Hansa last names	901	from Hansa mayors and Hansetag attendees
% of trips under Hansa last name	3.5	
% of trips outside Hansa network by captains with Hansa last names	3.8	
% of trips within Hansa network by captains with Hansa last names	5	
# of decades	20	190 years collapsed into 20 decades[a]
# of non-zero bilateral connections	17,648	per decade, balanced panel
# of non-zero bilateral connections with value observation	17,383	per decade, balanced panel
# of observations in bilateral sample	9,712,320	balanced panel
% of zeros in bilateral sample	99[b]	
# of bilateral connections	485,616	per decade, including unfrequented connections
Mean # of ships, bilateral	0.11	per bilateral connection and decade
Mean # of ships, non-zero	15.9	median at 2, 90th percentile at 29, among non-zero entries
Mean value, bilateral	2.5 Daler	per bilateral connection and decade
Mean value, non-zero	660 Daler	median at 39, 90th percentile at 736, among non-zero entries
# of trilateral connections	113,437	per decade
# of observations in trilateral sample	2,268,740	balanced panel
% of zeros in trilateral sample	91	
# of trilateral connections by last name	460,487	per decade
Mean # of ships, trilateral	0.35	per trilateral connection and decade
Mean # of ships, non-zero	4	median at 1, 90th percentile at 6 among non-zero entries
Mean value, trilateral	18.8 Daler	per trilateral connection and decade

Notes: City and trade descriptives in the restricted sample used throughout the paper.

[a]The first and the last one do not count ten years. We aggregate to decades rather than using the annual data due to computational limitations encountered when adding the dimension of a captain's home town below.

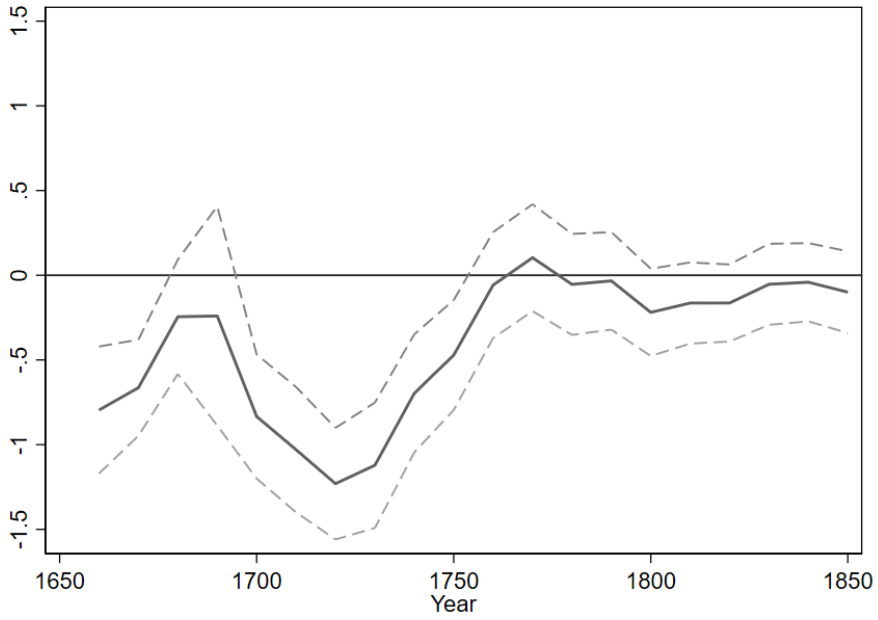
[b]This compares to roughly 80 per cent zeros in Jacks et al. (2020) who study the Empire's trade enhancing effect in British bilateral export data at the country and commodity level. As our data are city-level and we have far more cities than Jacks et al. (2020) has commodities, a higher share of zeros is to be expected.

Table 2.8.2: Hanseatic captains continued to relatively dominate their old trade routes

	Share of a home town A on a trade route B to C					
	(1)	(2)	(3)	(4)	(5)	(6)
Home town share 1668-1679	-0.0154***	-0.0156***	-0.0264***			
	(0.00101)	(0.00138)	(0.00101)			
Home town share 1668-1679 x All Hansa	0.00902***	0.00832**	0.00480			
	(0.00298)	(0.00357)	(0.00303)			
Highest Share 1668-1679				-0.0178***	-0.0152***	-0.0140***
				(0.00109)	(0.00129)	(0.000955)
Highest Share 1668-1679 x All Hansa				0.00984***	0.00847***	0.00474*
				(0.00303)	(0.00327)	(0.00264)
Fixed Effects:						
– Origin x Time	✓	✓		✓	✓	
– Destination x Time	✓			✓		
– Destination x Home town x Time		✓			✓	
– Origin x Destination x Time			✓			✓
– Home town x Time	✓		✓	✓		✓
– All Hansa x Time	✓	✓	✓	✓	✓	✓
Observations	419,201	313,233	419,167	973,866	777,909	967,261

Notes: Standard errors in parentheses. * p <0.10, ** p <0.05, *** p <0.01. The dependent variable is the share of a home town on a trade route in the decades 1680 to 1850. The independent variables in columns 1-3 are that home town's share on the trade route in the initial decade, 1668-1679, and the interaction of this share with a dummy for Hansa captains on former Hansa routes. The independent variables in columns 4-6 are a dummy indicating whether that home town had the highest share on a trade route in the initial decade, 1668-1679, and the interaction of this dummy with a dummy for Hansa captains on former Hansa routes. Across specifications, the first coefficient therefore measures how having dominated trade initially predicts trade shares in later decades and the second adds how this differs for Hansa captains on former Hansa routes. Estimation via OLS.

Figure 2.8.2: Aggregate bilateral results

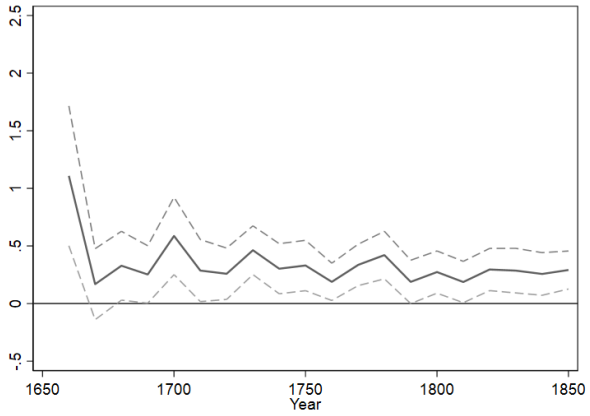


Notes: Aggregate bilateral estimation on decadal data. Estimation of equation 2.5.1 using PPML and a complete set of decade-interacted origin and destination fixed effects. Shown are point estimates of β_t accompanied by 95% confidence intervals. Standard errors clustered on city pairs. Mean of flows: 0.11. This regression result is presented in detail in Appendix Table B.B.4.

Figure 2.8.3: Hometown-level results



(a) Trilateral Hansa dummy, (origin x destination x time) and (home town x time) fixed effects. Identifying as Hanseatic by captain's home town.



(b) Trilateral Hansa dummy, (origin x destination x time) and (home town x time) fixed effects. Identifying as Hanseatic by last name.

Notes: PPML estimation including same harbour dummies (equal to one when either origin or destination coincide with the home town) of equation 2.6.1 with (origin x destination x time) and (home town x time) fixed effects in Figure 2.8.3a and of equation 2.6.1 with (origin x destination x time) and (last name x time) fixed effects in Figure 2.8.3b. Shown are point estimates of β_t accompanied by 95% confidence intervals. Standard errors clustered on city pairs. The regression result for Figure 2.8.3a is presented in detail in Appendix Table B.B.6.

Chapter 3

AI and Services-Led Growth: Evidence from Indian Job Adverts

This is the latest working paper version. The paper is currently a Revise & Resubmit at the *Journal of Human Resources*.[1]

Alexander Copestake[2], Max Marczinek[3], Ashley Pople[4], Katherine Stapleton[5]

Abstract. We document near-exponential growth in the demand for artificial intelligence (AI)-related skills in India's services sector from 2015, using a new dataset of online vacancies from its largest jobs website. We evaluate the impact of demand for AI skills on establishment-level non-AI postings, using a shift-share design that exploits variation in exposure to new AI inventions. We find that AI adoption significantly reduces growth in non-AI posting volumes and wage offers, particularly for highly skilled managerial and professional occupations, non-routine work, and analytical and communication tasks.

JEL Classification Codes: J23, O33

Keywords: Artificial Intelligence, Labor Markets, Wages, Services, Development

[1]First version: May 2021. This document is an output from the research initiative 'Structural Transformation and Economic Growth' (STEG), a programme funded by the Foreign, Commonwealth & Development Office (FCDO). We also gratefully acknowledge financial support by the Oxford Centre for the Study of African Economies (CSAE). The views expressed herein are those of the authors and should not be attributed to the FCDO or any of the institutions with which the authors are affiliated. Declarations of interest: none.

[2]International Monetary Fund, acopestake@imf.org.

[3]University of Oxford, max.marczinek@economics.ox.ac.uk. Nuffield College, New Road, OX1 1NF, Oxford, UK.

[4]University of Oxford, ashley.pople@economics.ox.ac.uk.

[5]World Bank, kstapleton@worldbank.org.

3.1 Introduction

Rapid advances in artificial intelligence (AI) have spurred an intense debate about its labor market consequences.[6] Online job adverts show that demand for AI-related skills grew almost exponentially and concurrently in several countries around the world from 2015 to 2019 (Figure 3.1.1).[7,8] Yet detailed empirical evidence on the extent of AI deployment and its distributional impacts in developing countries remains scarce. The use cases and impacts of AI need not be the same in developing countries as in advanced economies, and AI could have important implications for development pathways. Consider two examples: machine translation (between languages) and machine transcription (from handwriting to digital text). Improvements in the former could reduce barriers to exporting services and increase employment in developing countries (Brynjolfsson et al. 2019, Baldwin & Forslid 2020), while improvements in the latter could incentivize outsourcing firms in developing countries to shrink their workforce through automation. Understanding the net impact of such technologies on employment opportunities is therefore important for developing countries considering whether to promote a services-led development model.

In this paper, we shed light on the labor market impacts of AI in India – the archetypical pioneer of a services-led development model. We investigate these effects in the urban, white-collar services sector using a new dataset of online job adverts posted from 2010 to 2019 on India's largest jobs platform, which hosts an estimated 60 percent of the country's online vacancies. Following a growing literature including Rock (2019) and Acemoglu et al. (2022), we use the demand for AI-related skills, observed in the text of posted job descriptions, as a proxy for AI deployment.

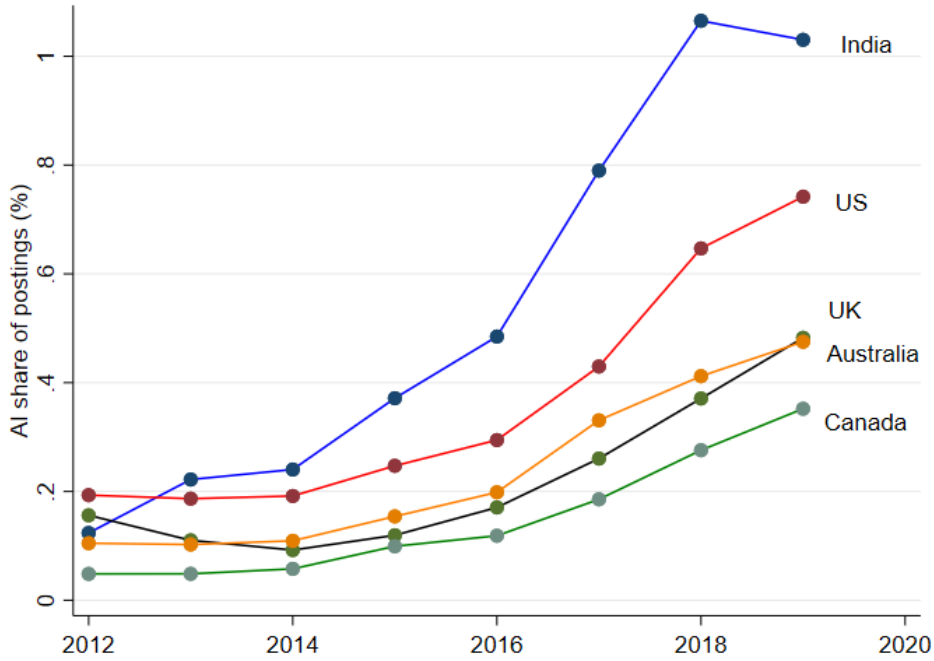
Given the scarcity of evidence on adoption of AI in developing countries, we first provide four stylized facts on the demand for AI-related skills (hereafter 'AI demand') in India using the job adverts data. First, we observe a rapid take-off in the rate of AI demand after 2015, rising

[6]To fix definitions, we consider artificial intelligence (AI) 'the theory and development of computer systems able to perform tasks normally requiring human intelligence, such as visual perception, speech recognition, decision-making, and translation between languages' (Oxford English Dictionary 2020). Machine learning, the sub-field responsible for many of the recent commercial applications of AI, comprises 'the statistical techniques that enable computers and algorithms to learn, predict and perform tasks from large amounts of data without being explicitly programmed' (Acemoglu & Restrepo 2019). We henceforth use 'AI' as an umbrella term encompassing machine learning.

[7]Demand for AI-related skills has continued to grow, especially following the launch of ChatGPT in 2022 and subsequent advances in and applications of generative AI. In this paper, we focus on the previous wave of innovation in the 2010s, based on the combination of deep neural networks and graphics processing units (for an early example, see Krizhevsky et al. 2012). Applications of this technology include recommendation engines (e.g., embedded in television or music streaming services), image recognition (e.g., automated photo tagging), fraud detection, risk scoring, spam filtering, and machine translation. For a detailed discussion, see Klinger et al. (2018).

[8]While online job postings in different countries and from different data providers may not be directly comparable, we see a sharp and simultaneous increase in AI-related postings in all cases.

Figure 3.1.1: AI share of online job postings, by country



Notes: This graph shows the share of all online vacancies – across all sectors – that specify particular AI skills, with these skills defined as described in Section 3.2.2. Data for India is that used in this paper; data for all other countries is from Lightcast, which does not cover India.

from 0.37 percent of all job postings that year to 1.03 percent in 2019, with take-up particularly pronounced in the IT, finance and business process outsourcing (BPO) industries. This growth coincides with the take-off in developed countries and reflects an increase in demand for specific 'deep learning' skills. Demand for AI skills is largely driven by technical roles such as software development, but there is also substantial demand for AI skills in generalist and managerial roles. Second, AI roles are heavily concentrated in the largest firms and a few key technology clusters – particularly Bengaluru, Mumbai, Hyderabad, Pune, Chennai and Delhi – similar to patterns observed in the USA by Babina et al. (2024) and McElheran et al. (2024). Third, adverts for AI roles tend to feature more verbs associated with complex, creative, and data-driven tasks. Fourth, AI roles tend to require substantially more education, particularly graduate degrees, while also paying significantly higher wages. Even after controlling for a host of fixed effects, posts demanding AI skills still pay a 13-17 percent salary premium, which is similar to the 11 percent estimate found in the USA (Alekseeva et al. 2020).

We then investigate the impact of this AI adoption on demand for non-AI labor at the establishment level. In theory, the impact is ambiguous. Advances in machine learning have been conceptualized as reducing the cost, or improving the quality, of the task of 'prediction', which is

prevalent in many occupations (Agrawal et al. 2018).[9] While this suggests displacement of labor, such improvements could also expand labor demand by reducing costs or increasing the quality of production and hence raising productivity (Webb 2020, Acemoglu et al. 2022, Acemoglu 2024). In addition, AI could create new tasks, spawn further innovation, or incentivize changes in organisational structure (Brynjolfsson et al. 2017, Cockburn et al. 2018, Klinger et al. 2018, Goldfarb et al. 2023, Agrawal et al. 2024). Finally, AI could reduce barriers to services trade, potentially expanding the set of tasks that are outsourced (Brynjolfsson et al. 2019, Baldwin & Forslid 2020).

To measure the net impact empirically, we use a long difference and shift-share instrumental variables strategy, examining how AI adoption relates to growth in AI and non-AI job postings and wage offers between 2010-12 and 2017-19. AI adoption is measured as the year in which an establishment (defined as a firm-city pair) first posts an AI vacancy. We focus on mature 'incumbent' establishments that posted on the platform at both baseline and endline, allowing us to observe their labor demand before and after the mid-decade surge in AI adoption. To focus on AI usage rather than AI production, we exclude AI-producing industries, as in Acemoglu et al. (2022).[10] We exploit plausibly exogenous variation in firms' *ex ante* exposure to advances in AI to account for unobservable differences between AI and non-AI adopters. Specifically, we first measure the extent to which the workforce in 2010-12 within establishments performed tasks later automatable by AI, using the occupation-level metric developed by Webb (2020) that assesses overlap between job tasks and patented AI capabilities. We then create a shift-share instrument for AI adoption by combining these occupational AI 'shocks' with establishment-level occupation vacancy shares at baseline. This instrument assumes that, within cells of similar establishments formed by our granular fixed effects, those more exposed to AI-automatable tasks are more likely to adopt AI, independent of other characteristics. Our first stage results confirm higher AI adoption in more exposed establishments. We present *ex ante* arguments (as described in Borusyak, Hull & Jaravel 2021) and *ex post* robustness tests (following Goldsmith-Pinkham et al. 2020) in support of our instrument's validity.

Overall, we find a significant negative impact of AI adoption on establishment-level labor demand. A one percentage point increase in the probability of an establishment adopting AI between 2010-12 and 2017-19 – which corresponds to approximately one standard deviation higher

[9]For example, a back office employee of a multinational bank takes the scrawled handwriting on a mortgage application form as a visual input, then generates a typed name of the applicant as predicted output.

[10]Specifically, we drop IT services and the education sector, reflecting that these industries produce the vast majority of AI patents and AI research (Klinger et al. 2020, Chahal et al. 2021).

ex ante exposure – results in an 8.1 percentage point reduction in the growth of non-AI vacancy postings. The effect on total vacancies (including AI vacancies) is similarly substantial, with a 7.8 percentage point decrease. The small difference between these two numbers indicates that the growth in vacancies constituted by new AI posts is far outweighed by the displacement effect in the much larger set of non-AI vacancies. These effects are also substantial when compared to the median growth in postings of 24.9 percent over the same period.

To better understand the mechanisms driving these effects and identify which types of jobs and tasks are most impacted by AI adoption, we analyze the heterogeneous impacts across occupations, skill requirements, and tasks. These negative impacts primarily reflect a reduction in demand for skilled managerial and professional occupations, non-routine work, and analytical and communication tasks. First, examining the impact on establishment-level vacancy composition by occupation group, we find a substantial reduction in higher-skilled professional and managerial occupations, notably engineering professionals and general and corporate managers. Second, using the classification of Acemoglu & Autor (2011), we find that AI demand lowers demand for occupations that are typically non-routine task intensive. This contrasts with previous recent waves of technological change that primarily lowered demand for routine tasks. Importantly, this negative impact on non-routine task-intensive occupations persists within the affected managerial and professional occupation groupings. We find similar negative impacts on 'abstract' task intensity, as defined in Autor & Dorn (2013), both within and between occupation groups. Finally, we leverage the richness of our text data to generate a more granular and flexible measure of task content following Michaels et al. (2018). We count verbs in job descriptions and classify them according to meaning using Roget's Thesaurus, and find that AI adoption reduces demand for verbs related to 'intellectual faculties'. In particular, there is a reduction in the frequency of verbs related to: 'precursory conditions' (such as investigate, scrutinize, research, explore, examine); 'extension of thought' (such as predict, forecast, anticipate, memorize, recall); and those related to 'means of communicating ideas' (such as narrate or describe). Such verbs align closely with the capabilities of new machine learning algorithms, suggesting that AI technologies are reducing the need for human labor in tasks involving analysis, prediction, and complex communication.

How does this displacement affect wage offers for new hires? We estimate that a one percentage point increase in the probability of an establishment adopting AI between 2010-12 and 2017-19 reduces the growth rate of median non-AI wage offers by 5.5 percentage points, and that of all wage offers by 5.3 percentage points. This is driven almost entirely by a composition effect

across occupations: when controlling for occupation shares, we do not see statistically significant reductions across the wage offer distribution. Our findings thus suggest that AI adoption reduced the availability of relatively high-wage managerial and professional occupations within incumbent Indian white-collar services firms.

Finally, we investigate wider effects of AI adoption by aggregating across establishments, first within firms and second within districts.[11] At the firm level, we again find significant negative impacts of AI adoption on labor demand and wage offers. However, when aggregating instead to the district level, we do not detect a significant overall impact of AI adoption. A relatively small share of establishments adopted AI within our sample period; thus, it is likely that the impacts were not yet large enough to be observable at this broader geographic scale.

Taken together, our results suggest that the demand for AI-related skills has already had important effects on Indian service-sector establishments and firms, altering the distribution of labor demand and wage offers across occupations and tasks. Despite limited district-level effects at present, the wider economic implications of AI adoption are likely to grow as usage of the technology diffuses further throughout the economy.

Related literature: Our primary contribution is to the nascent literature on the impacts of AI on developing countries. To date, this has been primarily theoretical due to the scarcity of data on AI adoption in developing countries. Baldwin (2019) and Baldwin & Forslid (2020) have conjectured that machine learning, along with online platforms and software robots, could benefit developing countries by facilitating the offshoring of white-collar services. In contrast, Korinek & Stiglitz (2021) take an alternative view that developing countries will be negatively affected, because AI devalues their comparative advantage in lower-cost labor and natural resources. To the best of our knowledge, we are the first to use data on AI deployment in a developing country to provide causal evidence on such hypotheses.[12] Our evidence provides partial support for both perspectives. On one hand, we document rapid increases in labor demand for AI skills in India, particularly in the IT and financial services sectors, consistent with the upskilling of existing offshored white-collar services functions. On the other hand, we find a significant negative impact of AI deployment on non-AI labor demand and wages within establishments and firms, consistent

[11] Firms are national in our dataset, so this corresponds to collapsing respectively the cross-geography and cross-firm dimensions of our data.

[12] Notably, Brynjolfsson et al. (2019) provide evidence on the causal impact of AI on US-Latin America trade using evidence from AI deployment on eBay, a multinational e-commerce platform. Several other studies provide descriptive evidence on AI exposure, as opposed to observed AI adoption, using data on the cross-sectional distribution of occupations (e.g., Pizzinelli et al. 2023, Gmyrek et al. 2023). Alonso et al. (2022) consider the impact of automation technologies, broadly defined, on developing countries, using data on industrial robots.

with labor displacement. More broadly, our findings that these negative impacts are concentrated in relatively highly skilled managerial and professional occupations align with concerns about the potential for services-led growth to generate good jobs in a world where advances in AI enable the automation of an increasing share of service sector tasks.

We also contribute to the broader literatures on AI diffusion and its consequences. First, we offer new insights on the diffusion of AI skills in the important context of India. In contrast to evidence on slow diffusion within countries (e.g., Bloom et al. 2024), as well as theories of slow diffusion to low-income countries (e.g., Benhabib & Spiegel 2005), we find that growth in the demand for AI skills within India has been remarkably similar to, and even faster than, that documented in some advanced economies, highlighting the rapid global diffusion of AI capabilities. At the same time, we also find a high degree of clustering, with our finding that AI hiring is highly concentrated within a small number of firms and tech hubs aligning with patterns of diffusion in the US documented by Babina et al. (2024) and McElheran et al. (2024).

Turning to the impacts of AI, we find both similarities and differences to findings in advanced economies. Although India has a different labor market structure to high-income countries, our findings of negative within-establishment effects of AI on labor demand echo those of Grennan & Michaely (2020) and Acemoglu et al. (2022) in the US. Our findings that these negative effects are driven by higher-skilled occupations and analytical tasks are also consistent with Grennan & Michaely (2020) and Webb (2020), yet contrast with the results of Acemoglu et al. (2022) that AI exposure predicts *increased* demand for skill families relating to engineering, analysis, marketing, finance and IT in the US. Our findings thus highlight the potentially nuanced impacts of AI across contexts, with occupation-level impacts differing in India despite similar establishment-level displacement effects. Our evidence that AI negatively affects non-routine task-intensive occupations also stands in contrast to findings for previous waves of technological change, such as computerization, that lowered demand for routine task-intensive occupations (for instance, Autor et al. 2003, Goos & Manning 2007, Goos et al. 2014).[13]

Finally, we add granularity to the existing literature on the labor demand impacts of AI by leveraging the rich data available on our partner's platform. Building on Michaels et al. (2018), who study the historic occupational distribution in the US, we use the specific verbs in job descriptions to show how AI disproportionately displaces particular tasks, such as forecasting and prediction. Our dataset also includes detailed information on wage offers – in contrast to

[13]This finding that the pattern of impact of machine learning is very different to previous waves is also consistent with emerging evidence on the impacts of generative AI (e.g., Brynjolfsson et al. (2023), Noy & Zhang (2023)).

commonly used platforms in advanced economies, where wage offer information is missing for the majority of posts (Batra et al. 2023) – allowing us to investigate the implications of AI adoption for wage offers. In doing so, we also contribute to the growing literature that uses online vacancy postings to explore labor market effects (e.g. Deming & Kahn 2018, Adams et al. 2020, Javorcik et al. 2020, Babina et al. 2023, 2024).

The rest of this paper proceeds as follows. Section 3.2 introduces the data, and Section 3.3 presents detailed descriptives on AI demand in the Indian white-collar services sector. Section 3.4 lays out our empirical strategy, and Section 3.5 presents our main findings on the impact of AI adoption in hiring on non-AI labor demand and wage offers. Section 3.6 provides several robustness checks and extensions, including exploring effects beyond the establishment level. Section 3.7 concludes. The Appendix provides further detail on the construction of our dataset, as well as additional results and robustness checks.

3.2 Data

3.2.1 Vacancy data

Our primary dataset is a random sample of 80 percent of job vacancies posted on India's largest online job platform between 2010 and 2019.[14] Firms primarily advertise on the platform and conduct subsequent recruitment and hiring offline. The platform estimates a 60 percent market share of Indian online job vacancies in 2020. We focus on the white-collar services sector, for which the data is most representative of overall job vacancies, dropping posts from the manufacturing and agriculture sectors. Our dataset includes approximately 15.5 million service sector job postings, equating to an average of 1.5 million per year, skewed towards later years. This compares to an estimated 19 million people formally employed in the service sector in India in 2021, according to the Ministry of labor and Employment.[15]

Firms submitting job vacancies on the platform are required to use a standardized template. Hence, all posts include information on the job title, industry, role category, location, skills required, salary and experience ranges and educational requirements. The job postings also feature an open text section for the job description, exclusively in English as they are within the formal services sector. We manually map industries and occupations into the National Industrial Classification (NIC) at the two-digit level and National Classification of Occupations (NCO) at the four-digit level, covering 99 percent of all vacancies. We also harmonise city names and add

[14]The company requested to remain undisclosed.
[15]https://static.pib.gov.in/WriteReadData/specificdocs/documents/2022/jan/doc20221104101.pdf

geolocations, separating out overseas job postings. Using the geolocations, we match cities to districts using the 2011 census. We focus on full-time jobs, which make up 96 percent of the sample, and drop the small number of part-time and non-permanent positions.

Table 3.2.1 provides key descriptives for our main sample separately for all jobs and for AI jobs only. As with all causal analysis in the paper, we focus on 'AI-using' industries and only on establishments posting vacancies both in the baseline (2010-2012) and endline (2017-2019). While 36% of all vacancies across the baseline and endline were posted in the baseline, the same share is only 2.6% for AI jobs, reflecting their growth between these periods. AI vacancies are heavily concentrated, with the 5 most active establishments (firms) posting 63.8% (73%) of all AI jobs. While 21% of all vacancies are in Banking and Financial Services, 71% of AI vacancies are situated in this industry. Similarly, only 2% of overall vacancies are for computer programmers, but 16% of AI vacancies are for this occupation. Finally, AI vacancies require on average more work experience and pay substantially higher average wages.

Table 3.2.1: Key descriptives of vacancy data, all jobs and AI jobs only

	All Jobs	AI Jobs
# Job Posts	1,902,969	16,827
% Posts in Baseline (2010-2012)	36	2.6
# Establishments	24,485	944
% Posts by 5 Most Active Establishments	4.4	63.8
Mean # Posts Per Establishment	78	18
# Firms	5,605	525
% Posts by 5 Most Active Firms	10	73
# Cities	442	32
% Posts in Top 5 Cities	64	92
# Industries (NIC 2008 features 86)	29	24
% Posts in BPO and Call Centres	25	5
% Posts in Banking and Financial Services	21	71
# 4-digit Occupations (NCO 2004 features 96)	94	49
% Posts for Computer Programmers	2	16
% Posts for Business Professionals	22	12
% Posts for Accountants	7	2
Mean Required Work Experience, Years	1.86	2.2
Mean Annual Salary, Rupees	167,000	389,000

Notes: This table shows descriptives for the restricted sample, which includes only establishments that post vacancies in both the baseline (2010-2012) and endline (2017-2019), and only posts from these years, and only posts in service sector, AI-using (not producing) industries. These are thus the vacancy posts underlying our data set for the main regressions. The number of NIC 2008 industries in our data is smaller than the total number of these industries. These are 2-digit industries, called divisions in the NIC 2008. The 86 divisions include manufacturing, agriculture, and mining, whereas we only include the services sector.

Our dataset has several advantages relative to administrative datasets. First, and most importantly, we can directly observe demand for AI skills in the text of job descriptions. In contrast, national surveys during the period (specifically, the National Sample Survey in 2011-12

and the Periodic Labour Force Survey in 2017-18) only record broad occupational categories, which are insufficient to identify the use of AI skills.[16] Similarly, firm-level datasets such as Prowess only include variables commonly reported in periodic financial statements, which are insufficiently granular to identify specific AI-related investment. Second, our vacancy dataset has broader coverage, with more than two million services vacancies in 2018, from 40,000 unique firms, compared to 12,000 workers surveyed in the 2017-18 Periodic Labour Force Survey and 2,000 firms recorded in Prowess (see Appendix C.A.2 for details). Lastly, our data are at a higher frequency than the representative sample surveys, which only took place in 2011-12 and 2017-18 and so can provide only limited insight into employment dynamics in the Indian services sector around the rapid take-off in AI demand from 2015.[17]

However, the richness of the vacancy data comes with certain shortcomings, notably that online vacancies are not representative of all vacancies and only proxy for firm hiring behaviour. Broadly speaking, our vacancy data best represents urban, white-collar service sector jobs, with a greater representation of the IT & BPO, Finance, Insurance and Real Estate, Professional and Business Services industries relative to administrative datasets. We provide a detailed overview of the coverage and representativeness of the data in Appendix C.A.2, where we benchmark the vacancy data relative to nationally-representative labor surveys and firm-level data from Prowess. Vacancies are disproportionately concentrated in urban centres, but at least one post appears for nearly every district in India.[18] Although the share of formal jobs is lower in India than in advanced economies, we find a ratio of 0.08 of annual average job postings to total formal employment. This is similar to the ratio of 0.09 for annual US job postings from Burning Glass Technologies relative to total US employment from 2010-2018 (Acemoglu et al. 2022).

3.2.2 Measuring AI demand

Despite the prominence of AI in popular discussion, firm-level data on AI adoption remains scarce (Raj & Seamans 2018). In the absence of more granular data, a growing body of work uses technology-related human capital to proxy for technology adoption.[19] Human capital is a key input for deploying AI systems and the scarcity of skilled AI professionals is well recognized as a primary obstacle to widespread adoption of AI, with top-tier scientists earning extremely high

[16]For instance, the closest categories to a machine learning engineer in the National Occupational Classification are the broad codes '2132 – Computer Programmers' and '3122 – Computer Assistants'.

[17]Specifically, we receive daily data, which we aggregate to an annual frequency in our regressions.

[18]The full geographic distribution of posts is shown in Appendix Figure C.A.1.

[19]For example, Rock (2019) and Benzell et al. (2019) use LinkedIn profiles to construct firm-level measures of engineering and IT talent, while Harrigan et al. (2020) use the firm-level employment share of 'technology workers' in French matched worker-firm data as a measure of technology adoption. Trefler & Sun (2022) build a measure of AI deployment for mobile apps.

salaries and frequently being bought out of academic positions. A global survey of around 2,000 companies by McKinsey Global Institute (2019) found that the primary method for sourcing AI talent and capabilities was external hiring. Most companies build their AI capabilities in house rather than buying or licensing capabilities from large technology companies. Even when firms subcontract AI services, they typically require some in-house AI-related human capital to oversee and manage the process. We therefore follow this emerging literature in using the demand for AI skills as a proxy for the extent of AI adoption within firms, assuming the two are positively correlated.

Online job vacancy data are well suited to the measurement of demand for very specific technology-related human capital owing to the detailed text data on the skills demanded for specific roles. To measure firm demand for AI skills, we classify job postings based on the text in the job description or skills requirements. Our main classification is the 'narrow' measure employed by Acemoglu et al. (2022), which categorises a post as an AI vacancy if it includes any word from a list of specific AI terms.[20] By using this narrow measure of AI skills, we reduce measurement error, although our estimates of demand for AI skills are likely to be a lower bound of the true level of adoption.

3.2.3 Verb categorisation

To shed light on the tasks that accompany AI-related skills, we follow Michaels et al. (2018) in using verbs mentioned in the text of job adverts as a proxy for task demand. We use the same list of 1,665 English verbs and the meaning of verbs from Roget's Thesaurus, which classifies words according to their underlying concepts and meanings. The Thesaurus is organized into 6 classes and 38 sections. The 6 classes are: Abstract Relations (ideas such as number, order and time); Space (movement, shapes and sizes); Matter (the physical world and human perception of it); Intellect (the human mind); Volition (the human will); and Emotion, Religion, and Morality.

3.3 Descriptives

Given the dearth of evidence on the adoption of AI in developing countries, this section presents four descriptive findings from our job postings data, documenting the trends in and composition

[20]Specifically, a post is categorized as AI-related if any of the following terms appear in either the 'job description' or 'skills required' fields: Machine Learning, Computer Vision, Machine Vision, Deep Learning, Virtual Agents, Image Recognition, Natural Language Processing, Speech Recognition, Pattern Recognition, Object Recognition, Neural Networks, AI ChatBot, Supervized Learning, Text Mining, Support Vector Machines, Unsupervized Learning, Image Processing, Mahout, Recommender Systems, Support Vector Machines (SVM), Random Forests, Latent Semantic Analysis, Sentiment Analysis / Opinion Mining, Latent Dirichlet Allocation, Predictive Models, Kernel Methods, Keras, Gradient boosting, OpenCV, Xgboost, Libsvm, Word2Vec, Chatbot, Machine Translation and Sentiment Classification.

of posts listing AI skills in the white-collar services sector in India. Here, and throughout the paper, we focus on non-remote, full-time and permanent jobs, which make up 96 percent of the sample.

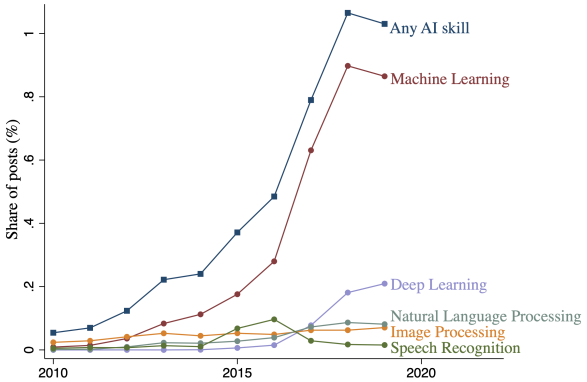
1. AI demand increased rapidly after 2015, with highest demand for technical data and management roles.

We first observe that the demand for AI skills increased rapidly after 2015, rising from 0.37 percent of all job postings that year to 1.03 percent in 2019 (Figure 3.3.1). This growth is largely driven by demand for general 'machine learning', with the sub-field 'deep learning' rising rapidly from relative obscurity to being the second most common AI term from 2017 onwards. AI demand grew steadily in the IT sector since 2011. However, after 2015, demand for AI-related skills also grew rapidly in 'AI-using' sectors. For instance, the financial sector saw a ten-fold growth in AI demand between 2016 and 2018, starting from a low base. The business process outsourcing and call centre sector saw a small boom in AI demand in earlier years, corresponding to a spike in demand for 'speech recognition'. The growth in AI demand in India coincides with the take-off in high-income countries, as shown in Figure 3.1.1.

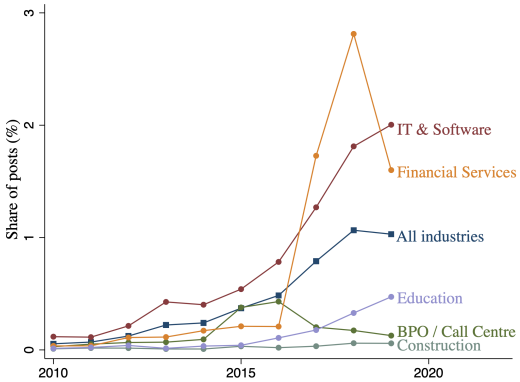
The most common AI role title by far is 'Software Developer', followed by other technical roles such as 'Data Analyst', 'Technical Lead' and 'Technical Architect' (Appendix Figure C.A.10). AI skills are also required in technical management roles, with titles such as 'Analytics Manager', 'VP - Analytics and BI', and 'Project Manager-IT/Software' also appearing in the top 20 AI-related roles. Yet there is also a long tail of more generalist roles, including 'Business Analyst', 'Trainee', 'Program Manager' and 'Product Manager'. Indeed, more than 25 percent of all AI vacancies are fragmented across other roles that each account for less than 1 percent of the total, highlighting that demand for AI skills is spread widely across many occupations.

Figure 3.3.1: Trends in AI demand

(a) Most demanded AI skills



(b) AI share of posts, by industry



Notes: Panel (a) shows the share of all vacancies that specify particular AI skills, for the top five most demanded skills. Appendix Figure C.A.11 decomposes overall demanded AI skills by establishment age. Panel (b) shows the share of vacancies that are AI vacancies, both for all industries together and within each of the top five industries by AI share. We present results on AI skills and posts by industry by the number of posts rather than their share in Appendix Figure C.A.6.

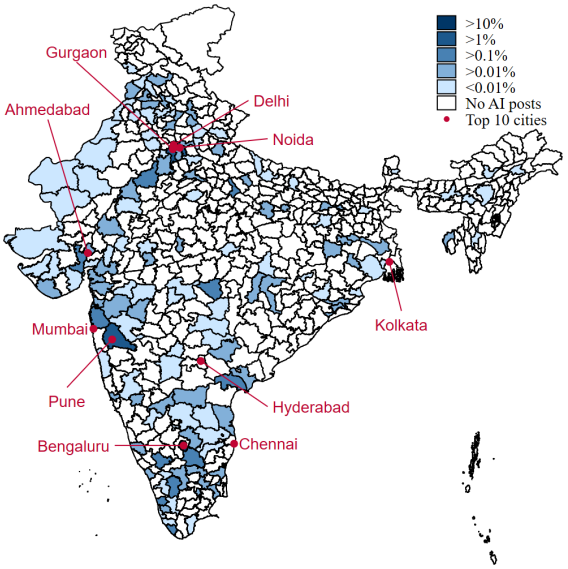
2. AI roles are highly concentrated in the largest firms and a few key technology clusters, particularly Bengaluru.

AI roles are highly concentrated in the largest firms. We proxy for firm size by the number of vacancies posted on the platform and compare the cumulative share of AI posts against the corresponding share of all posts. We find that the largest 14 firms are responsible for 10 percent of all vacancies, with each posting at least 50,000 vacancies. These account for 31 percent of all AI posts (Appendix Figure C.A.7). While there are some smaller firms that post a disproportionate number of AI posts, the largest AI-hiring firms are also the largest hirers in general.

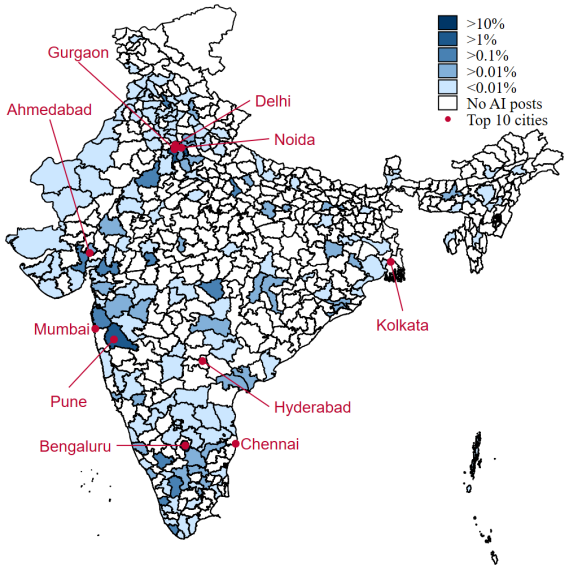
AI demand is also highly concentrated in large cities, particularly the major technology clusters around Bengaluru, Mumbai, Hyderabad and Delhi. Bengaluru alone has more than 30 percent of all AI vacancies across India (Appendix Figure C.A.8). Shares of AI demand in cities have been remarkably constant over the last decade, except for a prominent increase in AI activity in Mumbai as AI demand took off in the financial sector (Appendix Figure C.A.9).

Figure 3.3.2: Share of all AI posts by district, 2010-2019

(a) AI-using industries



(b) AI-producing industries



Notes: These maps shows the distribution of the share of all AI posts by particular districts, for the entire period 2010 to 2019. Panel (a) shows the distribution across AI-using industries, whereas panel (b) shows the distribution across AI-producing industries. Labels are shown for the top ten cities with the most AI posts. The majority of districts have few AI posts, since hiring is clustered in the largest cities.

3. AI job postings include more complex, creative, and data-driven tasks.

The text of AI job postings contains more verbs associated with complex, creative, and data-driven tasks. Extracting verbs from the job descriptions in AI and non-AI ads, we calculate the share of each verb relative to all verbs, and compare these shares in AI job posts to non-AI job posts. Table 3.3.1 shows the verbs with the largest frequency difference between AI and non-AI posts. Compared to non-AI posts, AI posts contain a higher share of complex tasks (e.g. learn and develop), creative tasks (e.g design and build), and data-driven tasks (e.g compute and predict). Appendix Table C.A.5 lists exemplary characteristic nouns to illustrate the context in which these verbs are used.

Table 3.3.1: Verbs that are more (less) common in AI posts

	Less common	**More common**
1	Manage	Learn
2	Account	Develop
3	Call	Process
4	Plan	Program
5	Maintain	Predict
6	Act	Build
7	Meet	Work
8	Recruit	Design
9	Sell	Advance
10	Report	Compute

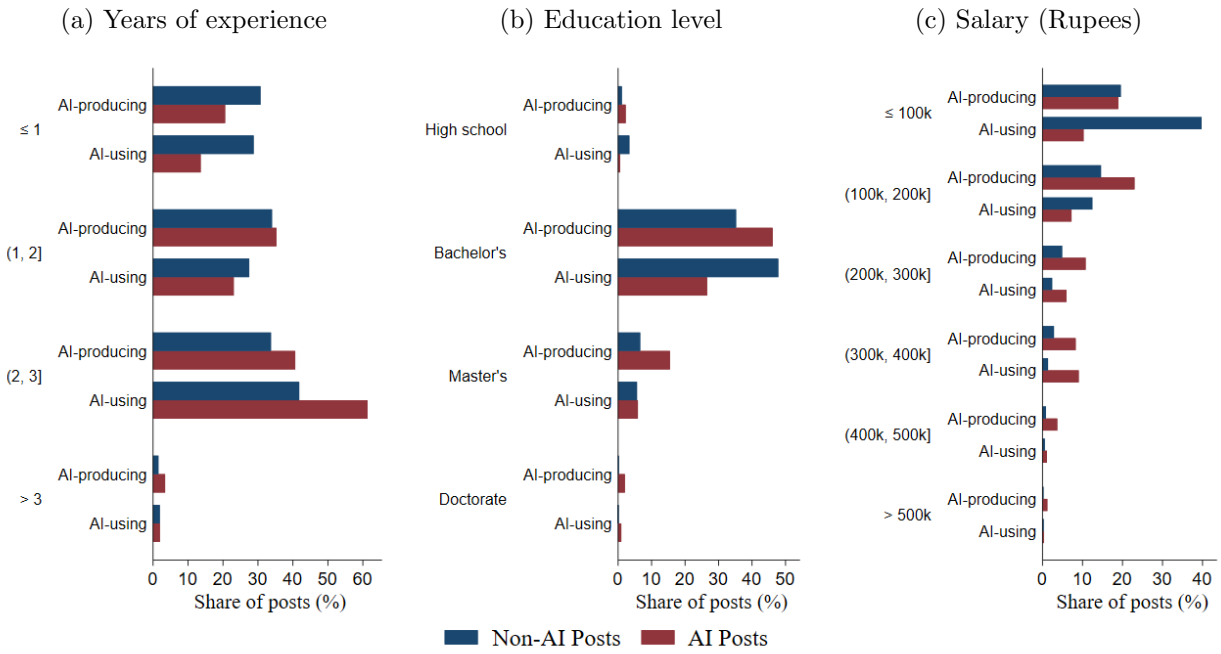
Notes: This table reports the ten verbs whose usage differs most between AI and non-AI job posts. First, for each verb in each establishment, we count the number of times that verb appears in all the AI posts by that establishment, and the number of times that verb appears in all the non-AI posts by that establishment. To form verb shares, we then compare these counts to the total number of verbs appearing across, respectively, all AI or all non-AI posts by that establishment. For each verb, we then estimate the difference in its appearance share between AI and non-AI posts using a regression with industry and region fixed effects and with standard errors clustered at the firm level. Positive (negative) differences imply that the corresponding verbs are more (less) likely to be included in AI posts. The ten verbs with significant positive coefficients of the largest magnitude are shown in the "More common" column. The ten verbs with significant negative coefficients of the largest magnitude are shown in the "Less common" column.

4. AI roles require more education and offer substantially higher wages than other white-collar services jobs.

AI roles list more advanced educational requirements and offer substantially higher wages than other white-collar services jobs advertized on the platform (Figure 3.3.3). AI vacancies are almost twice as likely as non-AI vacancies to require a master's degree, and more than seven times more likely to require a doctorate. They post a CPI-deflated median salary of ₹162,000 (approximately US$3,333, without adjusting for PPP), twice the median non-AI salary of ₹84,000 (US$1,666). This 'AI wage premium' remains high even when controlling for experience, education and firm fixed effects (20 percent), and occupation or role fixed effects (13-17 percent).[21]

[21]See Appendix Table C.A.4.

Figure 3.3.3: Hiring profile of AI vs. non-AI vacancies



Notes: These graphs compare the distribution of posts, for AI and non-AI vacancies, across experience, education and salary. This information is reported directly in the online jobs platform. Salaries are CPI deflated to 2010 levels. For experience and salary, the vacancy posts record a minimum and maximum value, so we take the midpoint of the specified range. AI posts are classified based on keywords, as described in Section 3.2.2.

In the rest of the paper, we build on these descriptive findings to assess the causal impact of AI adoption on labor demand. Our long-difference specification exploits the rapid increase in AI demand in the middle of our sample period (descriptive finding #1). Noting the concentration within particular cities and firms (descriptive finding #2), we do so using an establishment-level approach that controls for both city and firm-size fixed effects. Finally, given the specialization of AI roles in particular tasks (descriptive finding #3), and the 'AI wage premium' (descriptive finding #4), we examine the implications of AI adoption for the establishment-level composition of occupations and tasks and the wage offer distribution.

3.4 Empirical Strategy

3.4.1 Main specification

Our main specification assesses the impact of increased AI demand on changes in the demand for non-AI roles between our baseline period (2010 to 2012) and our endline period (2017 to 2019), spanning the take-off in AI demand from 2015.[22] We take a long-difference approach common in the automation and labor markets literature to focus on structural changes rather

[22]We pool within these periods in order to improve precision and maximise the probability that a firm advertises on the job postings platform during both time periods.

than short-term fluctuations in labor demand. Our primary unit of analysis is 'establishments', defined as firm-city pairs, as many firms report postings in several different cities. To focus on the effects of AI usage, rather than AI production, we exclude AI-producing industries as in Acemoglu et al. (2022).[23] Our main estimation sample contains almost 25,000 establishments posting approximately two million vacancies on the platform within our baseline and endline periods.[24] By definition, these establishments are part of relatively large, incumbent firms, as they existed already in 2010-12 and were still operational in 2017-19. We focus on this sub-set of all establishments to allow us to assess the medium-run *establishment-level* impacts of AI, recognizing that other channels (such as new AI-focused startups beginning in the mid-2010s) could also impact aggregate labor market outcomes.

Our main specification is

$$\Delta y_{fr,t-t_0} = \beta \cdot AdoptsAI_{fr,t-t_0} + \alpha_r + \alpha_i + \alpha_{f10} + \epsilon_{fr,t-t_0}, \tag{3.4.1}$$

where: $\Delta y_{fr,t-t_0}$ is the change in the natural logarithm of outcome Y_{frt} between 2010-2012 and 2017-19; $AdoptsAI_{fr,t-t_0}$ is a dummy taking value one for establishments that do not post an AI role in the baseline but do post an AI role in the endline; α_i and α_r are two-digit industry and city fixed effects; and α_{f10} is a firm size decile fixed effect, where firm size deciles are calculated over firm posting volumes in the baseline period.[25] We cluster standard errors at the firm level to account for common shocks across establishments within the same parent firm.

The primary identification concern is that AI adoption is systematically related to unobserved establishment-level factors that also increase non-AI labor demand, such as manager quality.[26] To address such potential endogeneity, we instrument AI adoption with 'AI exposure' to isolate changes resulting from supply-side technical advances in the capabilities of AI. We first take the occupation-level AI exposure measure developed by Webb (2020), which measures the extent to which workers' tasks can be performed by AI technologies using the degree of overlap between the text of AI patents and the text of O*NET job-task descriptions.[27] Occupations with a higher share of tasks that are capable of automation by AI are assigned a higher exposure value. We map

[23]Specifically, we drop education, IT, internet and e-commerce, telecom and internet service providers, which make up 34.8 percent of the sample.

[24]Appendix C.A.3 provides summary statistics for this sample.

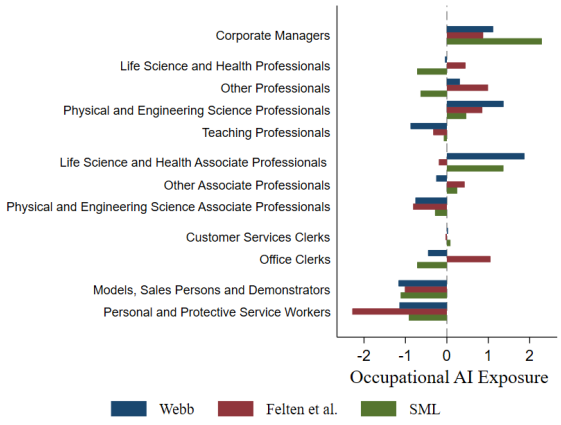
[25]The variables $\Delta y_{fr,t-t_0}$ approximate the growth in establishment outcomes between 2010-12 and 2017-19. Mathematically, for growth rate g defined by $Y_t = (1+g)Y_{t_0}$, and using the approximation that $\ln(1+g) \approx g$ for small g, we have $g = \ln Y_t - \ln Y_{t_0} = \Delta y_{t-t_0}$.

[26]For example, more innovative managers may be more likely to hire AI workers, but are also more productive and grow the business more quickly, hence also increasing non-AI labor demand. Appendix Table C.B.45 shows OLS results that support this suspected upward bias.

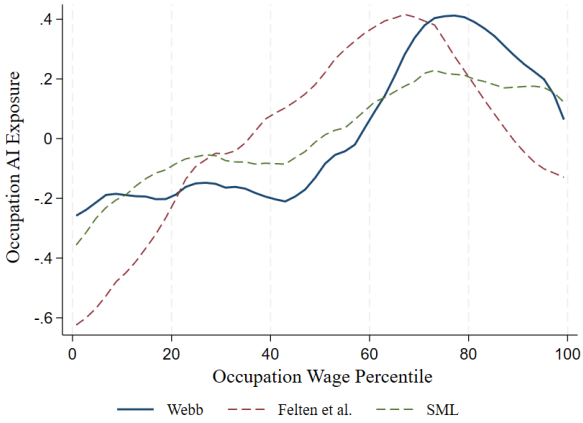
[27]These task descriptions are based on US occupations. While Indian occupations in general may have different task compositions, the white-collar service sector – which contains many multinational firms – is likely to be more similar. To the extent that this is not the case, it would also merely count against the strength of our first stage.

Figure 3.4.1: AI exposure by occupation group and by occupation wage offers

(a) AI exposure by occupation group



(b) AI exposure by occupation wage offers



Notes: These graphs display AI exposure by occupation group and wage offer. Panel (a) holds the average of the standardised AI exposure scores by broad occupation. The Figure in (b) shows a smoothed local polynomial regression of the Webb AI exposure measure on occupational wage offers. We first rank occupations by their average salary across all vacancy posts 2010-2019. We then plot the AI exposure associated with each, smoothing across a bandwidth of 10 percentage points. In addition to our main measure, from Webb (2020), we also show analogous results for the alternative measures (Felten et al. 2018, Mani et al. 2020) which we use in robustness checks in Appendix C.B.

the Webb (2020) exposure measure to the Indian National Classification of Occupations (NCO) 2004 at the four-digit level using publicly-available crosswalks. To capture establishment-wise exposure to AI-based automation, we then aggregate this measure to the establishment level by weighting across baseline establishment occupation shares. Specifically, we calculate:

$$Exposure_{fr,t_0} = \sum_o PostingShare_{fro}^{t_0} \cdot WebbExposure_o \tag{3.4.2}$$

where o represents occupations. We then standardize $Exposure_{fr,t_0}$ to have a mean of zero and a standard-deviation of one. Figure 3.4.1 shows the distribution of exposure scores across broad occupations and occupation-wise wage offer percentiles. AI exposure rises with wage offers up to a peak around the 80th percentile, before falling thereafter.[28]

We estimate the first stage:

$$AdoptsAI_{fr,t-t_0} = \gamma \cdot Exposure_{fr,t_0} + \alpha_r + \alpha_i + \alpha_{f10} + e_{fr,t-t_0}, \tag{3.4.3}$$

This is a linear probability model, such that the first stage coefficient γ estimates the increase in the probability of an establishment adopting AI between 2010-12 and 2017-19 that is associated with a one standard deviation rise in establishment exposure. When instrumenting equation 3.4.1 with equation 3.4.3, the coefficient β can thus be interpreted as the percentage point increase in the growth rate of outcome Y_{frt} that is associated with a 1 percent rise in the predicted

[28]Appendix Figure C.A.12 shows the distribution of AI exposure scores across industries.

probability of a firm adopting AI between 2010-12 and 2017-19.

We discuss several variations of this main specification in Section 3.6, including alternatives to a linear probability model in the first stage (Appendix Tables C.B.42 and C.B.43) and alternatives to using an AI adoption dummy (Appendix Tables C.B.29 to C.B.32).

3.4.2 Identification assumptions

Our case for a causal interpretation stems from the baseline posting 'shares' by occupation in 2010-12 ($PostingShare_{fro}^{t_0}$) being exogenous with respect to the outcome variables (primarily changes in establishment-level postings between 2010-12 and 2017-19) after conditioning on our fixed effects. Our Bartik instrument combines these shares with the 'shocks' of occupation-varying measures of exposure to AI patents, $WebbExposure_o$.

Instrument relevance requires that the instrument must correlate with the endogenous variable, $\Delta Adoption_{fr,t-t_0}$, conditional on region, industry, and firm size decile fixed effects, $\alpha_r, \alpha_i, \alpha_{f10}$. We find that AI exposure does indeed predict AI demand (Figure 3.4.2 Panel (a)). A one standard deviation higher establishment AI exposure score is associated with a significant 1.06 percent increase ($p<0.01$) in the probability of an establishment adopting AI between 2010-12 and 2017-19, after controlling for region, industry and firm size decile fixed effects.[29] Panel (b) of Figure 3.4.2 shows the relationship between AI exposure and AI adoption over time and confirms that the differential appeared at the same time that machine learning techniques became widely used. For instance, the AI share of vacancies posted by the most exposed quintile of establishments was relatively similar to other exposure quintiles until 2016, before rapidly diverging to reach almost 8 percent in 2018.

Instrument validity requires that, conditional on region, industry and firm size decile fixed effects, our AI exposure instrument is exogenous with respect to the error term. We require:

$$\mathbb{E}\Big[Exposure_{fr,t_0} \cdot \epsilon_{fr,t-t_0} | \alpha_r, \alpha_i, \alpha_{f10}\Big] = 0. \tag{3.4.4}$$

This assumption has a more intuitive interpretation. Goldsmith-Pinkham et al. (2020) highlight that this exogenous shares approach is equivalent to combining many difference-in-differences designs. Conditional on fixed effects, we need to argue that each of these is valid. For each occupation o, we require:

$$Cov\Big[\epsilon_{fr,t-t_0}, PostingShare_{fro}^{t_0} | \alpha_r, \alpha_i, \alpha_{f10}\Big] = 0. \tag{3.4.5}$$

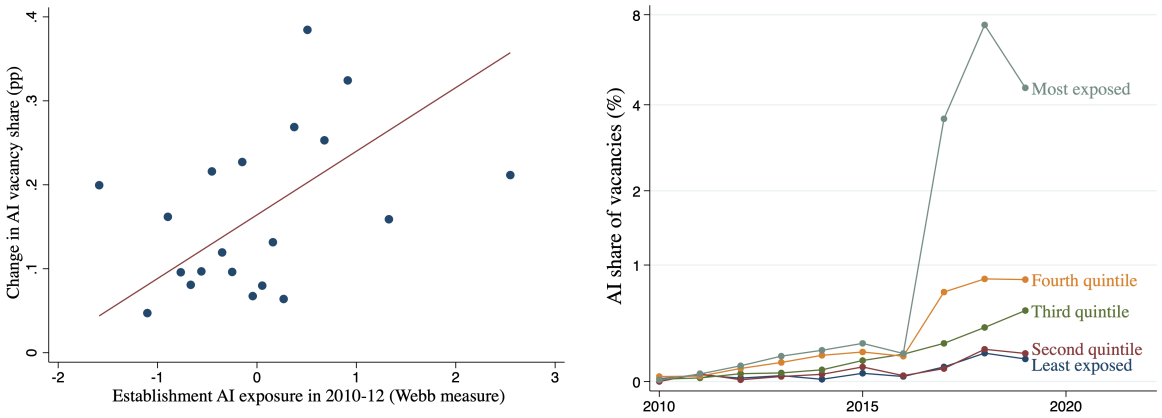
As our outcome is measured in first differences, this is a parallel trends assumption. Thus, we

[29] Full results are provided in Appendix Table C.B.4.

Figure 3.4.2: Impact of AI exposure on the AI share of establishments' posts

(a) Long differenced AI share vs. exposure

(b) Annual AI share by exposure quintile



Notes: These graphs show the relationship between AI exposure and the AI share of establishments' posts. The binned scatter plot in (a) summarizes the relationship between baseline AI exposure and establishments' change in AI vacancy share between 2010-12 and 2017-19, after partialling out region, industry and firm-decile fixed effects following Cattaneo et al. (2023). Panel (b) plots the time variation in this relationship, using an inverse hyperbolic sine scale for the y-axis.

claim that establishments with different baseline posting shares would have had parallel trends in labor demand and wage offers in the absence of the occupation-level AI shock. While this exogeneity assumption cannot be directly tested, we suggest both *ex-ante* arguments and *ex post* robustness tests.

First, we argue that, conditional on fixed effects, baseline occupation shares are exogenous to future trends in labor demand and wage offers. $\epsilon_{fr,t-t_0}$ is a residual trend in labor demand and wage offers between the baseline (2010-12) and endline (2017-19) periods. Baseline occupation shares, $PostingShare_{fro}^{t_0}$, are predetermined by definition. Thus, for these shares to be influenced by $\epsilon_{fr,t-t_0}$, establishments would have had to predict the AI boom and adjust their baseline occupation shares preemptively. We contend that this is highly unlikely, since the AI boom was sudden and unexpected. The boom was driven by significant advances in computer science, such as the AlexNet breakthrough in late 2012, which were rapidly commercialized and disseminated, for instance through the open-source release of TensorFlow in 2015. We argue that these breakthroughs were not predictable during the baseline period and thus should not have influenced the occupation structure of establishments at that time.

Could $\epsilon_{fr,t-t_0}$ be correlated with baseline posting shares for other reasons? While we cannot rule out all potential unobserved shocks affecting both baseline posting shares and subsequent labor demand trends, our fixed effects already account for varying trends by region, industry and firm size.[30] Moreover, the rapid adoption of machine learning techniques, as documented in

[30]For instance, the impact of increased AI deployment by firms in developed countries over our sample period, which could affect the extent to which they outsource back-office tasks to BPO firms in India,

Section 3.3, represent a large, generalized shock to the services sector, with no contemporaneous parallels of which we are aware. This limits the potential for our estimates to be biased by other unobserved technology shocks.

Following Goldsmith-Pinkham et al. (2020), we provide three robustness checks to affirm that our instrument is valid, which are discussed in more detail in Appendix C.B.1. First, we investigate the correlates of the shares, finding that the instrument does not appear to be correlated with baseline controls. Second, we test for pre-trends by investigating whether the baseline occupation shares predict year-on-year growth in employment or wages, finding no predictive power. Finally, we compare a range of estimators and run over-identification tests, finding similar results across estimators. All three tests thus support the validity of our instrument.

3.4.3 Other potential identification concerns

Our approach also addresses two other potential identification concerns. First, it mitigates the risk of mean reversion bias when relating total labor demand to AI exposure. In a reduced-form setting, an idiosyncratic rise in hiring of AI-exposed workers during the baseline period could artificially inflate an establishment's measured exposure and subsequently reduce the total number of posts as the effect dissipates. This could lead to a spurious negative relationship between AI exposure and vacancy growth. However, our specification uses AI adoption between 2010-12 and 2017-19 as the main regressor, with AI exposure an instrument. For the same bias to occur, the idiosyncratic increase in 2010-12 posting of AI-exposed occupations would need to be systematically related to a higher probability of adopting AI between 2010-12 and 2017-19 – which is not the case for a truly idiosyncratic increase. Our approach in contrast captures the true effect: high baseline susceptibility to AI capabilities increases the probability of adopting AI in subsequent years, which in turn lowers total labor demand.

A second potential concern is that the hiring patterns of Indian firms may have influenced AI innovation, biasing our results. Building on our previous discussion of identification assumptions, we offer two additional observations. First, India was not a significant producer of new AI research during our study period, lagging behind major hubs like the USA and China, despite strengths in applied computer science (Perrault et al. 2019). Thus, global advances in AI patenting are unlikely to be affected by hiring patterns in Indian firms.[31] Second, by excluding establishments in AI-producing sectors and focusing on AI usage, we drop the establishments

would be absorbed by our industry fixed effects.

[31] Indeed, this means our shocks could plausibly be exogenous as well.

where such concerns would be most relevant.

3.5 Main Results

3.5.1 Impacts of AI on labor demand

We first examine the effects of AI adoption on non-AI vacancies. Table 3.5.1 shows the impact of AI adoption on the growth of non-AI vacancies and total vacancies, instrumenting with the Webb (2020) AI exposure measure. AI adoption reduces the growth in non-AI demand: a one percent increase in the predicted probability of adopting AI (which corresponds to approximately a one-standard deviation increase in AI exposure in the first stage) results in a 8.1 percentage point decrease ($p<0.01$) in non-AI vacancy growth at the establishment level between 2010-12 and 2017-19 in our main specification with region, industry and firm-decile fixed effects. There is a similarly sized decrease of 7.8 percentage points in the growth of total vacancies, highlighting that AI vacancies crowd out other white-collar services-sector vacancies.[32] Considering that the median growth rate in total and non-AI vacancies is 24.9 percent, these are substantial effects.

Table 3.5.1: Second stage: Impact of AI adoption on establishment non-AI vacancies

	Growth in Non-AI Vacancies			Growth in Total Vacancies		
	(1)	(2)	(3)	(4)	(5)	(6)
Adoption of AI	-7.975***	-12.90***	-8.064***	-7.737***	-12.47***	-7.840***
	(2.350)	(3.092)	(2.282)	(2.245)	(2.959)	(2.181)
Fixed Effects:						
– Region	✓	✓	✓	✓	✓	✓
– Industry	✓		✓	✓		✓
– Firm Decile		✓	✓		✓	✓
First Stage F-Stat	43.7	41.58	45.43	44.06	41.83	45.62
Observations	22,244	22,244	22,244	22,251	22,251	22,251

Notes: Standard errors in parentheses. * p <0.10, ** p <0.05, *** p <0.01. Standard errors clustered at the firm level. The independent variable is a dummy, which equals 1 for establishments that did not post AI vacancies in the baseline period, but posted AI vacancies in the endline period. The dependent variables are the change in the log of the respective establishment-level outcomes, specifically the number of non-AI vacancies and the total number of vacancies. AI adoption is instrumented by establishment AI exposure. Each coefficient therefore represents the percentage point impact upon the outcome variable of a one percent rise in the predicted probability of adopting AI. AI exposure is calculated as the standardized average of occupation AI exposure (from Webb 2020), over the occupations for which the establishment posts vacancies in 2010-12, weighted by the number of vacancies posted per occupation, as in Acemoglu et al. (2022).

[32] In both these results, the standard errors increase to 1.38 when implementing the adjustment by Lee et al. (2022), leaving the finding significant ($p<0.05$).

3.5.2 Mechanisms and distributional impacts

To understand the drivers of these effects and identify which types of jobs and tasks are most impacted by AI adoption, we explore the heterogeneous impacts of AI adoption across occupations, skill requirements and tasks. We find the following results:

AI adoption reduces growth in demand for higher-skilled occupations. We first study the effects of AI on postings at the occupation level, using India's NCO2004 classification of one-digit and two-digit occupations. We regress AI adoption on the change in the share of non-AI vacancies within each occupation group. Table 3.5.2 shows that the decline in overall demand is accompanied by a shift away from higher-skilled occupations: the categories of 'Professionals' and 'Managers' suffer large reductions in their respective shares. For every one percent increase in the predicted probability of adopting AI, we observe a 10.6 percentage point decrease in the growth of non-AI vacancies for 'Managers' and a 3.6 percentage point decrease for 'Professionals'. These impacts are significant, given that these occupations comprised 18 percent and 48 percent of all postings in the 2010-12 baseline period respectively. Conversely, we note increases in shares for lower-skilled workers, such as 'Personal, Sales and Security', 'Clerks' and 'Associate Professionals'. These findings align with Webb (2020) who found that AI disproportionately affects high-skilled jobs involving pattern detection, judgement and optimization, such as clinical laboratory technicians, chemical engineers and optometrists.

Disaggregating this result further within these two groups using two-digit occupation codes, Appendix Table C.B.15 shows that increased AI demand reduces the share of non-AI vacancies for 'Engineering Professionals', 'General Managers', and particularly strongly for 'Corporate Managers'. In the baseline period, 24 percent of all postings are for 'Engineering Professionals', 6 percent for 'General Managers', and 12 percent for 'Corporate Managers'. Reduced vacancies growth for these three occupations consequently plays a large role in our aggregate results.

AI reduces growth in demand for non-routine task-intensive occupations. We next aim to understand more about the relationship between impacts on occupations and their task content, following the literature examining routine and non-routine task intensity of occupations (e.g. Autor et al. (2003) and Acemoglu & Autor (2011)). We map the measures of Acemoglu & Autor (2011) onto India's NCO2004 classification of occupations and standardize them.[33] We then analyze the impact of instrumented AI demand on the change in logged routine and non-routine scores at the establishment-level. Table 3.5.3 documents the significant negative

[33]To allow us to consider percentage impacts using logs, we transform the standardized values x to $x + x_2$, where x_2 is the second lowest observed value, such that all transformed values are positive.

Table 3.5.2: Second stage: Impact of AI adoption on establishment non-AI vacancy shares, by occupation group

	Change in Non-AI Vacancy Shares				
	Personal, Sales & Security	Clerks	Associate Professionals	Professionals	Managers
Adoption of AI	2.074***	1.324***	10.46***	-3.637***	-10.59***
	(0.385)	(0.272)	(1.718)	(0.717)	(1.709)
Fixed Effects:					
– Region	✓	✓	✓	✓	✓
– Industry	✓	✓	✓	✓	✓
– Firm Decile	✓	✓	✓	✓	✓
First Stage F-Stat	45.43	45.43	45.43	45.43	45.43
Observations	22,244	22,244	22,244	22,244	22,244

Notes: Standard errors in parentheses. * p <0.10, ** p <0.05, *** p <0.01. Standard errors clustered at the firm level. The independent variable is a dummy, which equals 1 for establishments that did not post AI vacancies in the baseline period, but posted AI vacancies in the endline period. The dependent variables are the change in occupation shares. Occupation groups are 1-digit occupations from the NCO04. Each coefficient therefore represents the impact upon the outcome variable of a one percent rise in the predicted probability of adopting AI. AI exposure is calculated as the standardized average of occupation AI exposure (from Webb 2020), over the occupations for which the establishment posts vacancies in 2010-12, weighted by the number of vacancies posted per occupation, as in Acemoglu et al. (2022).

effect of increased AI demand on establishment-level growth in non-routine task intensity. A one percent higher predicted probability of AI adoption leads to 3.3 percentage point lower growth in non-routine task intensity. In contrast, the intensity of routine tasks is not significantly affected.[34]

This result of declining demand for non-routine tasks could reflect changing task demand within occupations or a shift in demand between occupations. We hence also ask how AI adoption affects establishment-level non-routine task intensity within occupation groups, as shown in Table 3.5.4. We find that AI adoption also reduces growth in demand for non-routine tasks within higher-skilled occupations. We find the strongest reduction in non-routine task growth for the category of 'Managers'.[35]

AI reduces demand for analytical and complex communication tasks. The above measures of task intensity of occupations relied on time-invariant measures of the task content of occupations from O*NET. But AI could also impact the task content within occupations over time. We therefore adopt a further, more granular, approach by studying the verbs mentioned

[34]Appendix Table C.B.16 finds similar results for abstract and routine tasks following Autor & Dorn (2013), with a negative impact on abstract tasks and no discernible impact on routine tasks. Appendix Table C.B.17 finds similar results while controlling for occupational shares.

[35]The estimated effect differs significantly across occupation groups, with the hypothesis of equal coefficients rejected at the 1 percent level (p value=0.00). Differences are also generally statistically significant for pairwise comparisons at the 1 percent level, but insignificant for comparisons of 'Professionals' and 'Personal', 'Professionals' and 'Clerks', and 'Personal' and 'Clerks'. Appendix Table C.B.18 repeats Table 3.5.4 for abstract tasks following Autor & Dorn (2013), finding similar results of a decline in growth in demand for abstract tasks within professional and managerial occupation groups, with the latter impact especially large.

Table 3.5.3: Second stage: Impact of AI adoption on establishment routine and non-routine tasks

	Growth in Non-Routine Tasks			Growth in Routine Tasks		
	(1)	(2)	(3)	(4)	(5)	(6)
Adoption of AI	-3.395***	-3.281***	-3.258***	0.287	0.243	0.273
	(0.889)	(0.894)	(0.849)	(0.841)	(0.846)	(0.804)
Fixed Effects:						
– Region	✓	✓	✓	✓	✓	✓
– Industry	✓		✓	✓		✓
– Firm Decile		✓	✓		✓	✓
First Stage F-Stat	52.87	59.01	56.31	52.87	59.01	56.31
Observations	22,251	22,251	22,251	22,251	22,251	22,251

Notes: Standard errors in parentheses. * p <0.10, ** p <0.05, *** p <0.01. Standard errors clustered at the firm level. The independent variable is a dummy, which equals 1 for establishments that did not post AI vacancies in the baseline period, but posted AI vacancies in the endline period. We also control for the change in the log of total posts. The dependent variables are the change in the log of the respective establishment-level outcomes, specifically routine and non-routine tasks, where we have shifted the mean of the distributions of routine and non-routine scores to be able to take logarithms. We average standardised routine and non-routine O*NET task contents by occupation, and form establishments' routine and non-routine task demand by weighting occupations by their standardized routine and non-routine scores. AI adoption is instrumented by establishment AI exposure. Each coefficient therefore represents the percentage point impact upon the outcome variable of a one percent rise in the predicted probability of adopting AI. AI exposure is calculated as the standardized average of occupation AI exposure (from Webb 2020), over the occupations for which the establishment posts vacancies in 2010-12, weighted by the number of vacancies posted per occupation, as in Acemoglu et al. (2022). The reported F statistic is the cluster-robust Kleibergen-Paap Wald F-statistic.

Table 3.5.4: Second stage: Impact of AI adoption on establishment non-routine tasks, by occupation group

	Growth in Non-Routine Tasks				
	Personal, Sales & Security	Clerks	Associate Professionals	Professionals	Managers
Adoption of AI	0.113	-1.363**	-6.378***	-5.564***	-10.76***
	(0.462)	(0.552)	(1.646)	(1.612)	(2.052)
Fixed Effects:					
– Region	✓	✓	✓	✓	✓
– Industry	✓	✓	✓	✓	✓
– Firm Decile	✓	✓	✓	✓	✓
First Stage F-Stat	47.03	46.82	49.29	48.36	47.19
Observations	22,251	22,251	22,251	22,251	22,251

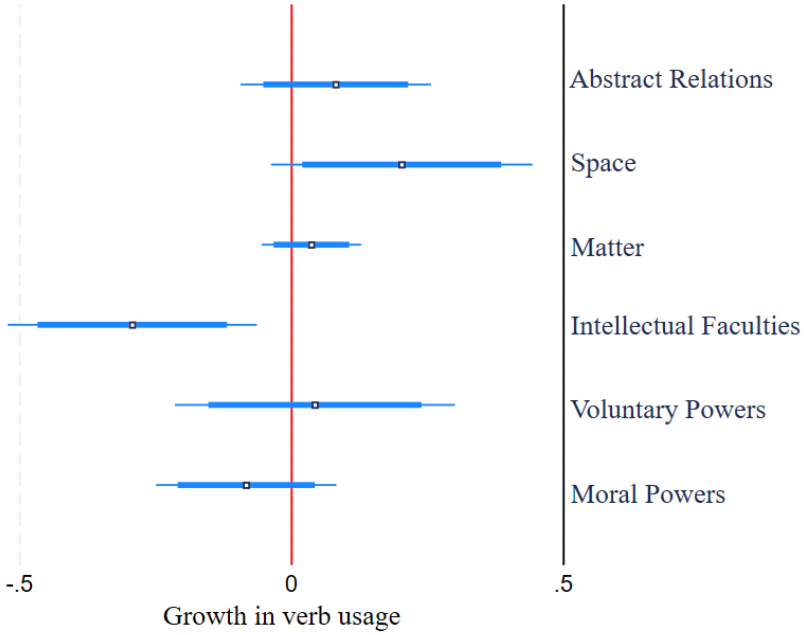
Notes: Standard errors in parentheses. * p <0.10, ** p <0.05, *** p <0.01. Standard errors clustered at the firm level. The independent variable is a dummy, which equals 1 for establishments that did not post AI vacancies in the baseline period, but posted AI vacancies in the endline period. The dependent variables are the change in the log of the respective establishment-level outcomes, specifically non-routine tasks, where we have shifted the mean of the distributions of routine and non-routine scores to be able to take logarithms. We average standardized routine and non-routine O*NET task contents by occupation, and form establishments' routine and non-routine task demand by weighting occupations by their standardized routine and non-routine scores. We control for the change in the respective occupation's vacancy share. AI adoption is instrumented by establishment AI exposure. Each coefficient therefore represents the percentage point impact upon the outcome variable of a one percent rise in the predicted probability of adopting AI. AI exposure is calculated as the standardized average of occupation AI exposure (from Webb 2020), over the occupations for which the establishment posts vacancies in 2010-12, weighted by the number of vacancies posted per occupation, as in Acemoglu et al. (2022). The reported F statistic is the cluster-robust Kleibergen-Paap Wald F-statistic.

within the text of job adverts. For each establishment, we count both the number of appearances of each verb and the total number of appearances of all verbs at baseline and endline. We thus construct a 'verb share' for each verb, as well as its change over time. We also repeat this procedure for groupings of verbs with similar meanings, using the verb classification of Roget's Thesaurus as discussed above.

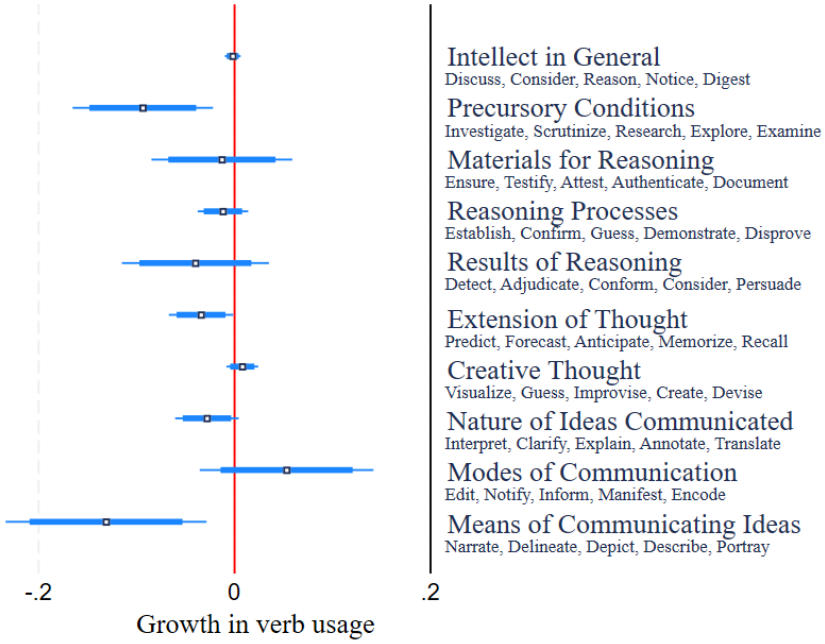
To analyze changing demand in tasks, we run regressions of a similar form to those in Table 3.5.1, using changes in verb shares as the dependent variable. Figure 3.5.1a shows that AI demand negatively impacts the share of verbs related to 'Intellectual Faculties' in Roget's Thesaurus. Specifically, a one percent increase in predicted AI adoption probability leads to a 0.29 percentage point decline in the share of those verbs. Figure 3.5.1b further breaks down this effect, showing statistically significantly negative effects on three sections within the 'Intellectual Faculties' category: i) 'Precursory Conditions', involving analytical tasks (e.g. 'investigate', 'research' and 'explore'), ii) 'Means of Communicating Ideas' involving complex communication tasts (e.g. 'narrate' and 'describe') and to a lesser extent iii) 'Extension of Thought', involving prediction (e.g. 'predict' and 'forecast'). These verbs align closely with known AI capabilities and support the idea that AI reduces the cost or improves the quality of the task of 'prediction' prevalent across many occupations (Agrawal et al. 2018).

Figure 3.5.1: Impact of 1 unit higher establishment AI demand growth on verb usage by class and section

(a) By verb class



(b) By verb section within Intellectual Faculties, and example verbs



Notes: These coefficient plots show the impact of AI adoption on verb share growth between 2010-2012 and 2017-2019, where verb shares are formed from counting verbs in job descriptions of job ads. Point estimates accompanied by 95 percent and 90 percent confidence intervals. Each coefficient is from a regression of type (3) in Table 3.5.1. Here, the outcome variable is growth in the share of verbs from the respective section or class. AI adoption is instrumented by establishment AI exposure. Each coefficient therefore represents the percentage point impact upon the outcome variable of a one percent rise in the predicted probability of adopting AI. AI exposure is calculated as the standardized average of occupation AI exposure (from Webb 2020), over the occupations for which the establishment posts vacancies in 2010-12, weighted by the number of vacancies posted per occupation, as in Acemoglu et al. (2022). Standard errors are clustered at the firm level, and we include region, firm size decile and industry fixed effects.

3.5.3 Impacts of AI on wage offers

We analyze wage offers by taking advantage of the platform's standardized job postings template, requiring firms to list wage offers for all vacancies.[36] Table 3.5.5 documents the impact of AI adoption on the growth of median wages for both non-AI postings and all job postings. We find that a one percent higher predicted probability of AI adoption reduces the growth of non-AI wage offers by 5.5 percentage points between 2010-12 and 2017-19, instrumenting with AI exposure and controlling for region, industry and firm size fixed effects. This negative effect remains largely unchanged when considering all posts, including AI postings, despite the higher wage offers associated with AI posts.

Table 3.5.5: Second stage: Impact of AI adoption on establishment non-AI wages

	Growth in Non-AI Median Wage			Growth in Overall Median Wage		
	(1)	(2)	(3)	(4)	(5)	(6)
Adoption of AI	-5.696***	-6.351***	-5.514***	-5.452***	-6.089***	-5.273***
	(1.485)	(1.630)	(1.423)	(1.424)	(1.566)	(1.366)
Fixed Effects:						
– Region	✓	✓	✓	✓	✓	✓
– Industry	✓		✓	✓		✓
– Firm Decile		✓	✓		✓	✓
First Stage F-Stat	45.14	42.91	46.86	45.46	43.12	47.02
Observations	22,064	22,064	22,064	22,071	22,071	22,071

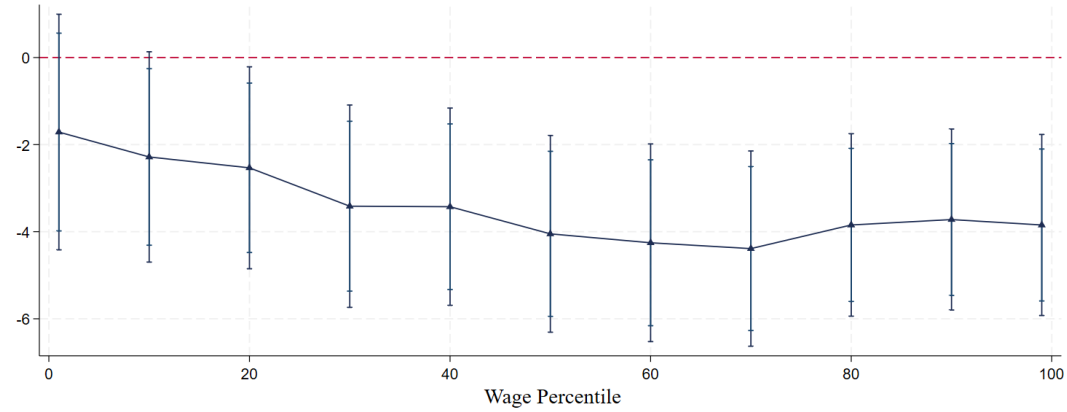
Notes: Standard errors in parentheses. * p <0.10, ** p <0.05, *** p <0.01. Standard errors clustered at the firm level. The independent variable is a dummy, which equals 1 for establishments that did not post AI vacancies in the baseline period, but posted AI vacancies in the endline period. The dependent variables are the change in the log of the respective establishment-level outcomes, specifically wages of non-AI vacancies and wages of all vacancies. AI adoption is instrumented by establishment AI exposure. Each coefficient therefore represents the percentage point impact upon the outcome variable of a one percent rise in the predicted probability of adopting AI. AI exposure is calculated as the standardized average of occupation AI exposure (from Webb 2020), over the occupations for which the establishment posts vacancies in 2010-12, weighted by the number of vacancies posted per occupation, as in Acemoglu et al. (2022).

The negative impact of AI demand on non-AI wage offers is evident across the wage offer distribution, with effects statistically significantly different from zero from the 20th percentile onwards. Figure 3.5.2 illustrates the impact of a one percent higher predicted probability of AI adoption on wage offer growth at various percentiles of the establishment-level distribution between 2010-12 and 2017-19, using the same AI exposure instrument and fixed effects as above. We observe a statistically significant reduction (at the 5 percent level) in establishment-level wage offers for non-AI jobs over time, ranging from -2.5 to -4.4 percent across the distribution.

The observed lower wage growth associated with AI demand could stem from two potential

[36]This contrasts with commonly used platforms in advanced economies, where wage offer information is missing for the majority of posts (Batra et al. 2023).

Figure 3.5.2: Impact of AI adoption on the wage offer distribution in non-AI posts

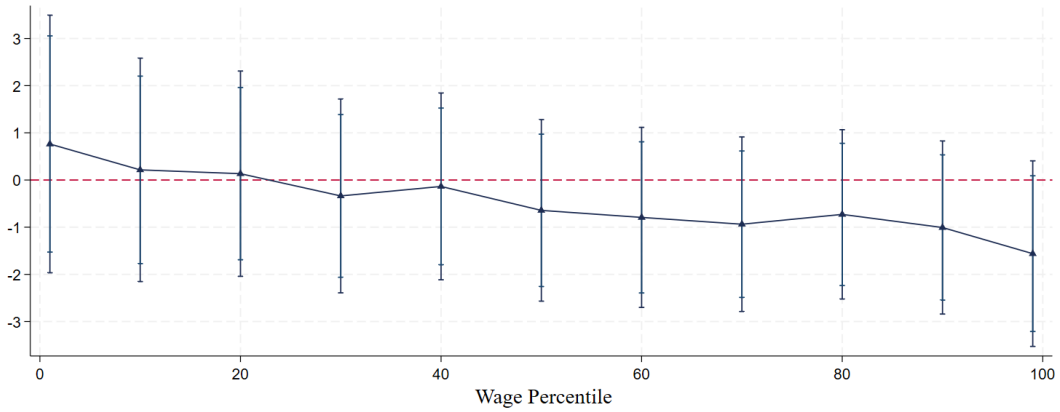


Notes: This coefficient plot shows the impact of a one percent rise in the predicted probability of adopting AI on wage growth over time across the distribution of establishment wage offers. As in Appendix Table C.B.20, AI adoption is instrumented by AI exposure. Each coefficient is from a regression of type (3) in Appendix Table C.B.20 and represents the impact of a one percent rise in the predicted probability of adopting AI on wage growth over time for a given percentile of the wage offer distribution. Standard errors are clustered at the firm level, and we include region, firm size decile and industry fixed effects. The outer (inner) error bars reflect 95 percent (90 percent) confidence intervals. Since AI posts make up only a small share of all roles in most establishments, the pattern is very similar across the distributions for all posts and for non-AI posts only.

sources: between-occupation or within-occupation effects. In other words, AI could change the occupational composition, thereby shifting the median wage offer, or it could reduce wage offers for individual occupations. Our findings in Table 3.5.2 showed that AI demand lowers hiring growth for the highest paid occupations, which likely contributes to the downward shift in overall wage offers. However, evaluating wage offer effects within occupation groups is challenging due to small sample sizes when splitting the data by occupation group.[37] When we add controls for changing occupation group shares, we find no significant declines in wage offer growth at the 5 percent level of significance, suggesting that composition effects are the primary driver of the results.

[37]We nevertheless report heterogeneous results on wage offers by occupation group in Appendix Table C.B.19, finding weak negative effects for higher-skilled occupation groups.

Figure 3.5.3: Impact of AI adoption on the non-AI wage offer distribution, holding occupational composition fixed



Notes: This coefficient plot shows the impact of a one percent rise in the predicted probability of adopting AI on wage growth over time across the distribution of establishment wage offers, controlling for the change in shares of 1-digit NCO04 occupations, leaving out 'Professionals'. As in Appendix Table C.B.20, AI adoption is instrumented by AI exposure. Each coefficient is from a regression of type (3) in Appendix Table C.B.20 and represents the impact of a one percent rise in the predicted probability of adopting AI on wage growth over time for a given percentile of the wage offer distribution. Standard errors are clustered at the firm level, and we include region, firm size decile and industry fixed effects. The outer (inner) error bars reflect 95 percent (90 percent) confidence intervals.

3.6 Robustness and Extensions

In this section, we first show that our findings are robust to several variations of our main specification and an event-study approach. We then investigate wider effects of AI, beyond the establishment level, and find significant impacts on (national) firms that adopt AI, but no robust evidence yet for aggregate district-level effects.

3.6.1 Alternative specifications

We show that our results are robust to several alternative specifications, with results provided in Appendix C.B. First, we show that our results are robust to adding additional controls for the baseline establishment posting shares of software engineers and sales and administrative professionals, following Acemoglu et al. (2022). By controlling for two broad occupations that see a decline due to computerization (Autor & Dorn 2013), we account for the possibility that firms experiencing increased demand for machine learning skills might also be software-engineering intensive or more affected by computerization. Our results remain robust with these additional controls (Tables C.B.21 to C.B.24).

Second, we show that our results are robust to using a shorter time period between baseline (2013-15) and endline (2017-19) in Tables C.B.25 to C.B.26. This approach draws on a larger sample of firms, as firms are more likely to post in this shorter timeframe, but at the cost of not spanning the entire takeoff in AI demand (Figure 3.3.1). Nonetheless, the results are similar

to those in the main specification. This shorter long-difference also allows us to differentiate between firms newly appearing in the data (termed 'startups' for convenience) and older firms ('incumbents') appearing in the data from the start.[38] In Tables C.B.27 and C.B.28, we show evidence of negative impacts on labor demand and wage offers in both groups of firms in the shorter long-difference specification, albeit less significant for newer firms where there is a weaker first stage.

Third, we demonstrate the robustness of our results using an alternative dependent variable that goes beyond our main 'AI demand' measure by taking into account the full quantity of AI vacancies posted by an establishment, rather than simply whether it is positive. In Tables C.B.29 to C.B.32, we replace $AdoptsAI_{fr,t-t0}$ in equation 3.4.1 with either the log of one plus the number of AI posts by an establishment or the inverse hyperbolic sine transformation of the same. These measures introduce an intensive margin (albeit at a cost to interpretability, as discussed in Chen & Roth (2023)). Nevertheless, the consistent findings across specifications provide reassurance that our results are not driven by our choice of independent variable.

Lastly, we confirm the robustness of our results to several other specification decisions. Our findings remain unchanged when including firm fixed effects (Table C.B.33) and when excluding the year 2019, in which a crisis in nonbank financial companies disrupted financial sector hiring (Tables C.B.34 and C.B.35; for details of the shock, see e.g., International Monetary Fund (2019, 2020), Agarwal (2024)). Our results are also unaffected by weighting by baseline establishment size (Tables C.B.36 and C.B.37), where this is proxied by the total number of vacancies posted in the baseline, with the upper 5 percent winsorized to reduce the dominance of the top firms as highlighted in Section 3.3. Likewise, our results go through when weighting vacancies to match occupation shares in labor force surveys (Tables C.B.38 and C.B.39) and when dropping occupations whose shares are sufficiently dissimilar (Tables C.B.40 and C.B.41). Our results are also robust to using a two-step consistent estimator accounting for endogenous treatment instead of a one-step linear probability model (Tables C.B.42 and C.B.43) following Wooldridge (2010). Our wage results are robust to using mean rather than median wage offers (Table C.B.44). Finally, when using the alternative AI exposure measure of Felten et al. (2018), we find similar negative impacts on wage offers and relative demand for professionals and managers (Appendix C.B.4).

[38]Specifically, we classify an establishment as an incumbent if it posts vacancies in the years 2010-2012 and 2017-19, and as a startup if it posts vacancies in 2013-15 and 2017-19 but not 2010-12.

3.6.2 Event-study approach

Our long-difference approach focuses on a fixed window (from 2010-12 to 2017-19) centered on the national (and global) take-off in AI demand in the middle of the 2010s. To provide an alternative perspective, we also take an event study approach, centering our analysis on the specific moment each establishment first posts an AI vacancy. This in turn allows the observation window to vary across establishments. In this case, the ideal experiment would compare the post-adoption outcomes of each AI-adopting establishment to those of an otherwise identical establishment that instead did not adopt AI. However, AI adopters are unlikely to resemble the average firm – for instance, they tend to be larger, as shown in Section 3.3.

To approach the ideal 'otherwise identical' counterfactual, we therefore follow Koch et al. (2021) in matching AI adopters to similar non-adopters using propensity scores. We construct these by running a Probit regression of AI adoption on a range of lagged establishment characteristics and hence deriving predicted adoption probabilities.[39] By accounting flexibly for a wide range of observables, this approach aims to control for unobservable differences between establishments. Rosenbaum & Rubin (1983) show that the propensity score is a balancing score, such that conditional on the propensity score, AI adoption is indeed orthogonal to observable establishment characteristics.

We then run an annual propensity score-weighted event-study regression at the level of establishments i:

$$Y_{it} = \alpha_i + \alpha_t + \sum_{k=-3\backslash -1}^{2} \beta_k \cdot 1(K_{it} = k) + \beta_{3+} \cdot 1(K_{it} \geq 3) + \epsilon_{it} \tag{3.6.1}$$

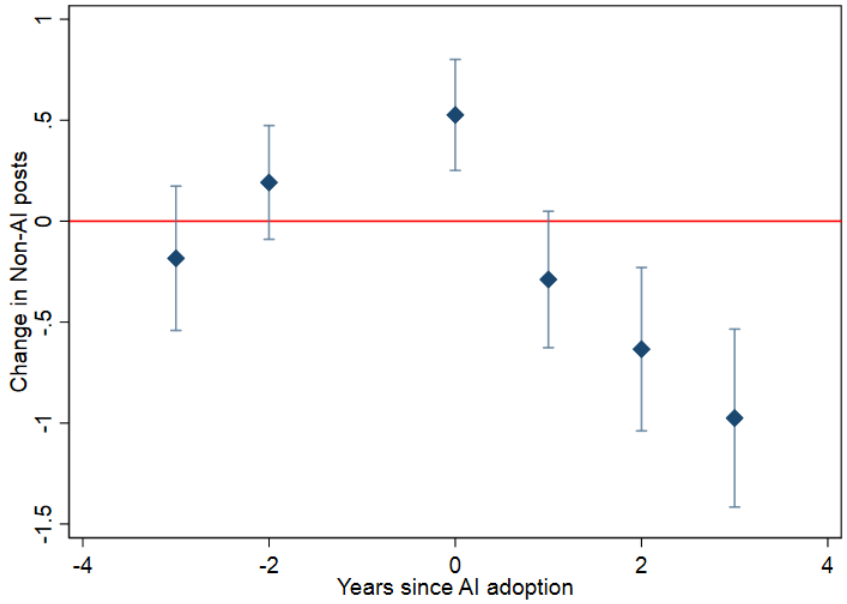
where Y_{it} is the inverse hyperbolic sine (IHS)-transformed number of job postings, α_i and α_t respectively are establishment and time fixed effects, and K_{it} is the time difference between the current year and adoption of AI.[40] We include three lags and leads, omitting the first lead to normalize the difference between adopters and similar non-adopters to zero in the year before adoption. The estimates β_k then measure the impact of AI adoption on the IHS of Y_{it}, relative to the level in similar non-adopters.

Figure 3.6.1 plots the resulting estimates. We find that, following an initial positive impact of AI adoption on non-AI vacancies, non-AI vacancies are significantly lower (at the 5 percent level) in the second year after adoption. This negative impact increases in magnitude in the

[39] Specifically, we include lags of firm size decile, establishment size, salary, experience levels, firm age, the standard deviation of salaries and experience, and several interaction terms, as shown in Appendix Table C.B.46.

[40] We take the inverse hyperbolic sine transformation to balance the panel such that firms that do not post within a year are not dropped from the sample.

Figure 3.6.1: Event study for non-AI postings following AI adoption



Notes: Two way fixed effects on a balanced panel. The outcome variable is the IHS-transformed number of establishment non-AI vacancies. Adopters are matched to never adopters by propensity score weighting, with propensity scores from a Probit regression of establishment characteristics on AI adoption (see Appendix Table C.B.46 for details). We leave out the first lead as the base period and cluster standard errors on establishments.

third year following AI adoption. Non-AI vacancies are 0.7 IHS points lower for adopters in the second year after adoption, and one IHS point lower three years after adoption.[41]

3.6.3 Wider effects

Does AI demand affect broader labor market outcomes, beyond the establishment level? First, we consider impacts at the firm level by aggregating nationally across a firm's constituent establishments. Tables 3.6.1 and 3.6.2 respectively show the results for the number of vacancies and the median wage offer. While the sample size drops substantially due to the aggregation and the F-statistic is correspondingly much weaker, we nonetheless find significant negative impacts of AI adoption on the growth in firms' overall posting of vacancies and their average wage offers.

Second, we consider impacts at the district level by aggregating across establishments in a given region (Table 3.6.3), with a significant reduction in sample size. We nonetheless find that AI exposure predicts AI adoption (column 1), but we do not find a significant impact on growth in non-AI or total vacancies (columns 2 and 3), and at most weak evidence ($p < 0.11$) for a negative effect on non-AI and overall wage offer growth (columns 4 and 5). Given the weak first stage in these long-difference regressions, we also explore results from a district-level

[41]These findings are also robust to using the imputation estimator of Borusyak, Jaravel & Spiess (2021).

event-study approach (Appendix Figure C.B.2), but again do not detect a significant impact of AI adoption on labor demand. Given that only 3.9 percent of establishments in our study used AI between 2010 and 2019, it is unsurprising that the impact of AI adoption on labor demand in AI-using industries is not detectable in district-level aggregates. This suggests that the effect, while present at the establishment level, was too small to influence broader labor market trends during our study period.[42]

Table 3.6.1: Second stage: Impact of AI adoption on firm non-AI vacancies

	Growth in Non-AI Vacancies			Growth in Total Vacancies		
	(1)	(2)	(3)	(4)	(5)	(6)
Adoption of AI	-4.603*	-12.05**	-3.140	-4.377**	-10.44**	-3.111*
	(2.445)	(5.934)	(2.079)	(2.186)	(4.722)	(1.860)
Fixed Effects:						
– Industry	✓		✓	✓		✓
– Firm Decile		✓	✓		✓	✓
First Stage F-Stat	18.38	7.268	18	19.7	8.475	19.16
Observations	5,701	5,702	5,701	5,703	5,704	5,703

Notes: Standard errors in parentheses. * p <0.10, ** p <0.05, *** p <0.01. Standard errors clustered at the firm level. The independent variable is a dummy, which equals 1 for establishments that did not post AI vacancies in the baseline period, but posted AI vacancies in the endline period. The dependent variables are the change in the log of the respective establishment-level outcomes, specifically the number of non-AI vacancies and the total number of vacancies. AI adoption is instrumented by establishment AI exposure. Each coefficient therefore represents the percentage point impact upon the outcome variable of a one percent rise in the predicted probability of adopting AI. The latter is instrumented by firm AI exposure. AI exposure is calculated as the standardized average of occupation AI exposure (from Webb 2020), over the occupations for which the firm posts vacancies in 2010-12, weighted by the number of vacancies posted per occupation, as in Acemoglu et al. (2022).

[42]Similarly, we do not see evidence of broader shifts in the sectoral composition of hiring over the period (Appendix Figure C.A.3).

Table 3.6.2: Second stage: Impact of AI adoption on firm non-AI wages

	Growth in Non-AI Median Wage			Growth in Overall Median Wage		
	(1)	(2)	(3)	(4)	(5)	(6)
Adoption of AI	-2.983**	-5.211**	-3.034**	-2.690**	-4.365**	-2.703**
	(1.204)	(2.507)	(1.228)	(1.089)	(2.044)	(1.105)
Fixed Effects:						
– Industry	✓		✓	✓		✓
– Firm Decile		✓	✓		✓	✓
First Stage F-Stat	18.65	7.356	18.23	19.97	8.565	19.38
Observations	5,671	5,672	5,671	5,673	5,674	5,673

Notes: Standard errors in parentheses. * p <0.10, ** p <0.05, *** p <0.01. Standard errors clustered at the firm level. The independent variable is a dummy, which equals 1 for establishments that did not post AI vacancies in the baseline period, but posted AI vacancies in the endline period. The dependent variables are the change in the log of the respective establishment-level outcomes, specifically wages of non-AI vacancies and wages of all vacancies. AI adoption is instrumented by establishment AI exposure. Each coefficient therefore represents the percentage point impact upon the outcome variable of a one percent rise in the predicted probability of adopting AI. The latter is instrumented by firm AI exposure. AI exposure is calculated as the standardized average of occupation AI exposure (from Webb 2020), over the occupations for which the firm posts vacancies in 2010-12, weighted by the number of vacancies posted per occupation, as in Acemoglu et al. (2022).

Table 3.6.3: AI adoption at the district level

	AI Adoption	Non-AI Vacancies	Vacancies	Non-AI Wages	Wages
	(1)	(2)	(3)	(4)	(5)
First stage:					
AI Exposure	0.0253**				
	(0.0107)				
Second stage:					
Adoption of AI		0.601	0.610	-2.430	-2.411
		(1.995)	(1.996)	(1.519)	(1.514)
First Stage F-Stat		5.569	5.569	6.300	6.300
Observations	399	399	399	399	399

Notes: Standard errors in parentheses. * p <0.10, ** p <0.05, *** p <0.01. Standard errors clustered at the district level. The dependent variables are the growth between 2010-12 and 2017-19 in district AI vacancies, non-AI vacancies, total vacancies, non-AI wages, and total wages, each in log differences. The independent variable for the first stage is district AI exposure, calculated as the standardized average of occupation AI exposure (from Webb 2020), over the occupations for which the district posts vacancies in 2010-12, weighted by the number of vacancies posted per occupation, as in Acemoglu et al. (2022). The first coefficient therefore represents the impact on the probability to adopt AI of a one-standard deviation rise in AI exposure. The independent variable for the second stage is the growth in district AI vacancies between 2010-12 and 2017-19, in log differences. The latter four coefficients therefore represent the percentage impact on the outcome variable of AI adoption, instrumented by district AI exposure.

3.7 Conclusion

AI could have important implications for a services-led development model, given the potential for recent advances in machine learning to automate many service sector occupations. In this paper, we use a new dataset of online vacancy posts from India's largest jobs website to shed light on the demand for, and implications of, AI capabilities in the white-collar services sector. We document a rapid take-off in demand for AI-related skills from 2015, particularly in the IT, finance and professional services industries, closely mirroring patterns found for advanced economies (Grennan & Michaely 2020, Acemoglu et al. 2022). AI jobs pay a substantial wage premium and are highly concentrated in certain industries, cities and large firms.

We assess the labor market effects of establishment-level demand for AI skills, as a proxy for AI deployment. Employing a long-difference shift-share specification based on exposure to patented advances in AI capabilities (Webb 2020), we find that AI adoption significantly reduces growth in non-AI postings and average wage offers. This decline primarily reflects a reduction in demand for skilled managerial and professional occupations, non-routine work, and analytical and communication tasks. This contrasts with previous findings on computerisation and robotics, which primarily affected routine tasks (Autor et al. 2003, Goos & Manning 2007, Goos et al. 2014).

Taking an even more granular approach and classifying verbs in the text of the job adverts using Roget's Thesaurus, we find that AI adoption reduces demand for verbs related to 'intellectual faculties', particularly those relating to investigation, prediction and description. These results are in line with the notion that machine learning reduces the cost or improves the quality of the task of 'prediction' (Agrawal et al. 2018).

Our findings underscore AI's potential to reduce the availability of high-skilled, high-wage white-collar jobs in developing countries — traditionally the target of services-led development strategies. However, while we observe significant negative impacts of AI adoption on labor demand and wage offers within firms, these effects do not translate to significant changes in district-level aggregates. This discrepancy likely stems from the relatively small proportion of establishments adopting AI during our sample period, such that their adoption decisions did not have a detectable impact on aggregate regional outcomes.

As AI adoption spreads, understanding the impacts on local labor markets in developing countries will be an increasingly important task for future research. The latest wave of innovation in generative AI, including large language models enabling applications like ChatGPT, has made this task yet more urgent (for a discussion of developing countries' exposure, see, for instance,

Cazzaniga et al. 2024). However, the collection of granular, representative data on usage of such technologies through official surveys remains at an early stage (Acemoglu et al. 2024), especially in developing countries. Alternative data sources such as job postings are thus a promising tool for answering these questions in such contexts.

Conclusion

This dissertation consists of three chapters in which I study how trade and labour markets absorb large shocks. The first followed early modern Baltic ports after the last Northern European outbreak of the plague, and showed that plagued regions emerged with a comparative advantage in capital-intensive goods that lasted nearly a century. The second traced the slow unravelling of the Hanseatic League. The chapter found that while the volume of trade on the League's old routes converged to what gravity would have predicted, the relationships running along those routes – carried by merchant families across generations of captains – did not. The third chapter moved to the present and examined how the diffusion of artificial intelligence is reshaping demand for skills in the Indian services sector.

The plague and the end of a major European trade network operate through factor proportions and through institutions, respectively; AI adoption in India operates through technology. Yet the three shocks are united in this dissertation in that I study how ports, regions, and firms adjust to them. Across chapters, I document a lasting economic response. Comparative advantage outlives demographic recovery in the plague chapter; commercial relationships outlive the institution that founded them in the Hanseatic chapter; and hiring patterns shift across skills and tasks in ways that descriptives alone would have missed in the AI chapter.

There is much that the three chapters do not do, and much that further work might. The plague analysis in chapter 1 identifies productivity growth from trade patterns and corroborates it with Swedish production records, but the supply side could be developed considerably further if firm- or workshop-level production data from other affected regions could be located and digitised. The Ricardian model could be enriched with a corresponding micro-foundation operating at the firm- or workshop-level. The chapter currently proposes such a micro-foundation; future work may provide empirical support for it and embed it in a general equilibrium structure. There is also a contemporary echo worth pursuing. Several advanced economies face simultaneous demographic contraction and stagnant productivity, and the conditions under which scarcity does or does not induce productivity growth – as it appears to have done in early eighteenth-century

Northern Europe – is a question with considerable policy weight today.

Chapter 2 raises questions that the data we have can identify but not fully answer. We show that family ties persist on the Hanseatic core routes long after the League itself has formally disappeared, but the data do not let us decisively separate competing explanations. There is also a question of generality. The Hanseatic League is one of several early modern trading networks; whether the patterns we document recur for the Italian republics, the Maghribi traders, or the Armenian diaspora networks would be a natural next step. Whether the patterns are still present in modern-day networks, for example among migrant communities or in developing country contexts, is perhaps an even more pertinent question.

The AI chapter is a snapshot of a still-developing process. We see displacement effects, but the vacancy data do not capture firm entry well, and the offsetting expansion – if there is one – in young firms or new sectors is something we can only speculate about. This issue is present in other ongoing and finished work on the economics of AI at the firm-level and we believe that future work will be able to address this puzzle. The international dimension also matters. We focus on AI adoption within India, but Indian services employment depend heavily on demand from advanced economies, and AI adoption abroad could affect that demand directly. Whether the domestic and foreign channels reinforce or offset one another is, at the time of writing, an open question. The concentration of AI hiring in a small number of firms and a small number of cities is a third issue worth pursuing; if the gains from this technology accrue narrowly while the displacement is broad, the distributional consequences could be severe.

Each of the chapters touches on questions that are alive in current policy debates. The plague chapter speaks to the broader question of how economies adjust to demographic decline, and suggests a silver lining. The Hanseatic chapter speaks to current concerns about the resilience of supply chains and the role of informal institutions in sustaining commerce when formal arrangements are disrupted. The AI chapter speaks directly to one of the most contested questions in contemporary labour economics, and does so for a country whose recent growth model puts it on the front line. While the shocks studied in this dissertation are of a different nature and centuries apart, I hope that my dissertation provides our discipline with a clearer perspective on how production, trade, and labour markets respond to such shocks.

Appendix A

Appendix for Chapter 1

Appendix A.A Data Appendix

Table A.A.1 holds key descriptives for the trade panels constructed from the Soundtoll records.

Table A.A.1: Descriptive Statistics

	Obs.	Mean	Std. Dev.	Median	Min	Max
Panel A: Bilateral Trade, Intensive Margin by Sector (origin × destination × sector × year)						
Exports (levels)	2,514,900	2.685	101.486	0.000	0.000	32110.792
Capital-intensive sector	2,514,900	0.400	0.490	0.000	0.000	1.000
Exports: capital-int. agriculture	502,980	4.101	78.956	0.000	0.000	8592.281
Exports: capital-int. manufacturing	502,980	1.588	45.716	0.000	0.000	7047.719
Exports: labour-int. agriculture	502,980	6.722	206.383	0.000	0.000	32110.792
Exports: labour-int. manufacturing	502,980	0.596	20.881	0.000	0.000	5183.438
Share of zeros (%)		97.7				
600 origins, 501 destinations, years 1668–1750						
Panel B: Bilateral Trade, Intensive Margin Overall (origin × destination × year)						
Exports (levels)	502,980	13.424	268.081	0.000	0.000	35175.146
Market share	502,980	0.013	0.099	0.000	0.000	1.000
Share of zeros (%)		95.1				
Panel C: Bilateral Trade, Extensive Margin (origin × destination × year)						
Number of exported goods	502,980	0.431	2.622	0.000	0	84
Any exports (dummy)	502,980	0.094	0.292	0.000	0	1
Panel D: Sectoral Productivity (origin × year)						
Log prod.: labour-int. agriculture	50,174	-1.237	5.128	-1.462	-39.675	34.770
Log prod.: capital-int. agriculture	50,184	0.560	6.425	-0.233	-47.423	31.429
Log prod.: labour-int. manufacturing	50,184	-0.452	4.165	0.000	-26.075	30.098
Log prod.: capital-int. manufacturing	50,184	5.105	8.704	2.020	-24.293	47.999
600 cities, years 1668–1750						
Panel E: City-Level Exports (origin × year)						
Total exports (levels)	49,800	135.585	1296.781	0.000	0.000	40149.656
Capital-intensive share	49,800	0.096	0.276	0.000	0.000	1.000
Zero total exports (dummy)	49,800	0.861	0.346	1.000	0.000	1.000
600 origins, years 1668–1750						

Notes: Summary statistics for the main datasets. Panel A is the bilateral origin-destination-sector-year panel. Panel B is the bilateral origin-destination-year panel used for market share regressions. Panel C summarises the extensive margin (number of goods exported per bilateral pair and year). Panel D reports log sectoral productivity estimates used in the productivity regressions. Panel E aggregates to the city-year level.

A.A.1 Commodities

The Soundtoll data do not provide researchers with cleaned information on commodities. Instead, the variable 'soort' denotes entries in words, specifically in the words of the custom official at the time. 143,855 separate descriptions are reported in Danish, which changed over the centuries. I use the Soundtoll Project's list of products translated into English and double-check in Danish etymological dictionaries. I manually clean goods as models trained on modern Danish are not suitable for Old Danish. Further, a lot of similar sounding words introduce complexity and require intuition, as closest neighbour matching can be highly misleading.

In 104,132 of 3.66 million shipments (2.8%, or 2.5% of value)[1], several goods are reported. I cannot discern which value accrues to either of the goods. Besides the relative uncommonness of this, there are two more reasons not to worry: for one, these two goods are usually within the same category. Also, these goods were regarded as substitutes by toll officials and carried the same duties, and only for this reason were reported together in one line rather than separately.

I code goods as granularly as I can. This granularity reflects the source's time period: many types of fish and textiles are recorded, whereas some shipments are mentioned to contain generic goods. I code individual types of fish as their own good; the compound category of fish contains all types that are not further specified. These differences in granularity matter little as I almost always aggregate trade up to the sectoral level.

For textiles, I distinguish raw materials (cotton, linen, wool) from processed textiles (band, gaze, ribbon, yarn, tissue, stockings), luxury textiles (flannel, satin, silk, velours, velvet) and naval textiles (sail/canvas, flag cloth). In cases where the data are insufficiently specific, I use compound categories (cloth, cotton textile, woollen textile). The compound categories distinguish, if possible, at least by cotton and wool. The good 'clothes' contains all types of fabric, textile, and clothes which are not further distinguishable, either by material or by degree of processing.

I arrive at 227 goods, 8 out of which only exist in the early trade data without bilateral coverage, and which therefore are never part of the analysis. In total, 143,855 goods descriptions are mapped onto 219 goods, so the average commodity appears in 9,000 variations. An example of different spellings is sturgeon, appearing in four variations: Stor, Større, Støre, and Stører. Another example of varieties of goods and spellings is iron, of which there are six variations: jerren, jeern, jern, jeren, osmundt jern, and osmunds jern, where the latter two denote specifically Swedish iron. Many variations in fact contain information on value or amounts: lead oxide appears in 15 variations, six of which specify value, e.g. 'gleide 3 ort' and 'gleide a 6 skilling'.

[1] Counting as double whenever 'og' (also) or 'etc' are contained in a goods description.

To complement my manual cleaning efforts with a simple robustness check, I used Claude's Opus 4.6 model to classify raw cargo descriptions.[2] For 84% of traded value, the algorithmic assignment is the same as the manual one. When accounting for a few inconsequential disagreements, I arrive at 96% overlap between the two approaches. The disagreements are between very closely related goods in the same sector, usually at a different aggregation, such as iron and iron works, planks and wood, cloth and sail, brandy and wine, goods and mercery, and meat and beef. Thus, the two approaches produce almost identical results.

From the manually cleaned goods, I then construct five sectors: labour-intensive agriculture, capital-intensive agriculture, labour-intensive manufacturing, capital-intensive manufacturing, and the remaining sector of unclassified goods. I distinguish agricultural from manufacturing goods based on their likely location of production, namely the countryside rather than the city. Capital-intensive goods require a high amount of capital that cannot be swapped for labour. This capital may consist of factories, furnaces, machinery, or tools necessary to produce the good. Labour-intensive goods are limited in the degree to which labour can be saved in their production. Labour-intensive agriculture is all of arable farming, whereas pastoral farming, mining, and processed foods including alcohols are capital-intensive agriculture. Labour-intensive manufacturing are largely textiles. Capital-intensive manufacturing produces goods such as processed metals, tools, and ship building materials. All 227 goods, accompanied by an assignment into the four sectors, are reported in Table A.A.3. 2% of traded value is in unclassified goods, 99% of which are generic goods and mercery. This implies that 98% of traded value is assigned into goods.

Table A.A.2 ranks the 30 most common goods traded between 1668 and 1750, which together make up 90% of the value of trade.

[2]The prompt was: "In this folder, you find the file "list from raw unique.dta". It contains a list of raw trade records, contained in the variable "soort", the value of the transaction ("totaldaler") and my manually cleaned commodity classificaton ("comm" and "good id"). Disregarding this manual classification, please make a new classification from scratch. A commodity that you can assign each entry of "soort" into has to be one of the commodities in "comm", however, or you can leave an entry as unclassified."

Table A.A.2: Top 30 goods by value

Good	Sector	Share of traded value	Cumulative share of traded value	Rank
Hemp	Labour-int. Agriculture	11.25	11.25	1
Rye	Labour-int. Agriculture	10.57	21.82	2
Salt	Capital-int. Agriculture	9.75	31.57	3
Wine	Capital-int. Agriculture	8.84	40.42	4
Wheat	Labour-int. Agriculture	8.63	49.05	5
Flax	Labour-int. Agriculture	6.1	55.15	6
Iron Works	Capital-int. Manufacturing	4.31	59.46	7
Iron	Capital-int. Agriculture	3.9	63.36	8
Tobacco	Labour-int. Agriculture	3.68	67.04	9
Planks	Capital-int. Manufacturing	3.22	70.27	10
Leather	Capital-int. Agriculture	2.81	73.07	11
Sugar	Labour-int. Agriculture	2.05	75.13	12
Unclassified	Unclassified	1.32	76.44	13
Ash	Capital-int. Manufacturing	1.22	77.67	14
Goods	Unclassified	1.22	78.88	15
Linnen	Labour-int. Manufacturing	1.14	80.03	16
Cloth	Labour-int. Manufacturing	1.12	81.14	17
Herring	Capital-int. Agriculture	1	82.14	18
Wax	Capital-int. Agriculture	.97	83.12	19
Brandy	Capital-int. Agriculture	.97	84.09	20
Barley	Labour-int. Agriculture	.8	84.89	21
Tow	Capital-int. Agriculture	.69	85.58	22
Wool	Capital-int. Agriculture	.68	86.26	23
Dye	Capital-int. Agriculture	.63	86.88	24
Tools	Capital-int. Manufacturing	.61	87.49	25
Silk	Labour-int. Manufacturing	.58	88.07	26
Wood	Labour-int. Agriculture	.56	88.62	27
Pepper	Labour-int. Agriculture	.55	89.17	28
Cotton	Labour-int. Agriculture	.48	89.65	29
Train Oil	Capital-int. Agriculture	.45	90.1	30

Notes: This table shows the top 30 goods by value of trade between 1668 and 1750.

Table A.A.3: Goods Classification

Animals	Capital-int. ag.	Rhine Wine	Capital-int. ag.	Ship Nails	Capital-int. man.	Tea	Labour-int. ag.
Arrack	Capital-int. ag.	Rum	Capital-int. ag.	Soda	Capital-int. man.	Tobacco	Labour-int. ag.
Beef	Capital-int. ag.	Salmon	Capital-int. ag.	Staves	Capital-int. man.	Trees	Labour-int. ag.
Beer	Capital-int. ag.	Salpeter	Capital-int. ag.	Steel	Capital-int. man.	Wheat	Labour-int. ag.
Bones	Capital-int. ag.	Salt	Capital-int. ag.	Swords	Capital-int. man.	Wood	Labour-int. ag.
Brandy	Capital-int. ag.	Seed	Capital-int. ag.	Tar	Capital-int. man.	Baize	Labour-int. man.
Bread	Capital-int. ag.	Sekt	Capital-int. ag.	Tiles	Capital-int. man.	Band	Labour-int. man.
Bristles	Capital-int. ag.	Soap	Capital-int. ag.	Tools	Capital-int. man.	Bay	Labour-int. man.
Buckling	Capital-int. ag.	Sprat	Capital-int. ag.	Turpentine	Capital-int. man.	Bowls	Labour-int. man.
Butter	Capital-int. ag.	Starch	Capital-int. ag.	Almonds	Labour-int. ag.	Chests	Labour-int. man.
Caviare	Capital-int. ag.	Stockfish	Capital-int. ag.	Amber	Labour-int. ag.	Cloth	Labour-int. man.
Chalk	Capital-int. ag.	Stones	Capital-int. ag.	Anise	Labour-int. ag.	Confectionaries	Labour-int. man.
Cheese	Capital-int. ag.	Sturgeon	Capital-int. ag.	Apples	Labour-int. ag.	Cotton Textile	Labour-int. man.
Cidre	Capital-int. ag.	Sulphur	Capital-int. ag.	Barley	Labour-int. ag.	Flag Cloth	Labour-int. man.
Coal	Capital-int. ag.	Talcum	Capital-int. ag.	Beans	Labour-int. ag.	Flannel	Labour-int. man.
Codfish	Capital-int. ag.	Tartar	Capital-int. ag.	Bushel	Labour-int. ag.	Folio	Labour-int. man.
Copper	Capital-int. ag.	Tin	Capital-int. ag.	Capers	Labour-int. ag.	Furniture	Labour-int. man.
Downs	Capital-int. ag.	Tow	Capital-int. ag.	Cardamon	Labour-int. ag.	Gaze	Labour-int. man.
Dye	Capital-int. ag.	Train Oil	Capital-int. ag.	Chestnuts	Labour-int. ag.	Gloves	Labour-int. man.
Eal	Capital-int. ag.	Treacle	Capital-int. ag.	Cinnamon	Labour-int. ag.	Hat	Labour-int. man.
Feather	Capital-int. ag.	Vetch	Capital-int. ag.	Cloves	Labour-int. ag.	Lace	Labour-int. man.
Fish	Capital-int. ag.	Vinegar	Capital-int. ag.	Cocoa	Labour-int. ag.	Lamps	Labour-int. man.
Flounders	Capital-int. ag.	Vitriol	Capital-int. ag.	Coffee	Labour-int. ag.	Linnen	Labour-int. man.
Flour	Capital-int. ag.	Wax	Capital-int. ag.	Cork	Labour-int. ag.	Mats	Labour-int. man.
Fur And Skin	Capital-int. ag.	Wine	Capital-int. ag.	Cotton	Labour-int. ag.	Paper	Labour-int. man.
Gin	Capital-int. ag.	Wool	Capital-int. ag.	Cucumbers	Labour-int. ag.	Ribbon	Labour-int. man.
Grass	Capital-int. ag.	Zinc	Capital-int. ag.	Dates	Labour-int. ag.	Ropes	Labour-int. man.
Gymsum	Capital-int. ag.	Alkali	Capital-int. man.	Dodder	Labour-int. ag.	Sail	Labour-int. man.
Haddock	Capital-int. ag.	Anchors	Capital-int. man.	Figs	Labour-int. ag.	Satin	Labour-int. man.
Hair	Capital-int. ag.	Ash	Capital-int. man.	Flax	Labour-int. ag.	Silk	Labour-int. man.
Ham	Capital-int. ag.	Barrels	Capital-int. man.	Frankincense	Labour-int. ag.	Spears	Labour-int. man.
Herring	Capital-int. ag.	Bells	Capital-int. man.	Fruit	Labour-int. ag.	Spokes	Labour-int. man.
Honey	Capital-int. ag.	Black Powder	Capital-int. man.	Ginger	Labour-int. ag.	Stockings	Labour-int. man.
Horn	Capital-int. ag.	Bottles	Capital-int. man.	Grapes	Labour-int. ag.	Tissue	Labour-int. man.
Horses And Horse Products	Capital-int. ag.	Bowsprits	Capital-int. man.	Groats	Labour-int. ag.	Trousers	Labour-int. man.
Iron	Capital-int. ag.	Brass	Capital-int. man.	Hemp And Hemp Products	Labour-int. ag.	Velours	Labour-int. man.
Lead	Capital-int. ag.	Brass Works	Capital-int. man.	Hops	Labour-int. ag.	Velvet	Labour-int. man.
Leadoxide	Capital-int. ag.	Bullets	Capital-int. man.	Ivory	Labour-int. ag.	Woollen Textile	Labour-int. man.
Leather	Capital-int. ag.	Cannons	Capital-int. man.	Juniper	Labour-int. ag.	Yarn	Labour-int. man.
Lings	Capital-int. ag.	Caps And Helmets	Capital-int. man.	Lemons	Labour-int. ag.	Ballast	Unclassified
Liquorice	Capital-int. ag.	Cement	Capital-int. man.	Linseed	Labour-int. ag.	Chinese Goods	Unclassified
Mackarel	Capital-int. ag.	China	Capital-int. man.	Millet	Labour-int. ag.	Comm	Unclassified
Malt	Capital-int. ag.	Glass	Capital-int. man.	Muscade	Labour-int. ag.	English Goods	Unclassified
Marble	Capital-int. ag.	Glue	Capital-int. man.	Nuts	Labour-int. ag.	Faroer Goods	Unclassified
Mead	Capital-int. ag.	Hoes And Shovles	Capital-int. man.	Oats	Labour-int. ag.	French Goods	Unclassified
Meat	Capital-int. ag.	Iron Works	Capital-int. man.	Olives	Labour-int. ag.	Goods	Unclassified
Mineral	Capital-int. ag.	Masts	Capital-int. man.	Peas	Labour-int. ag.	Icelandic Goods	Unclassified
Mustard	Capital-int. ag.	Medicine	Capital-int. man.	Pepper	Labour-int. ag.	Luggage	Unclassified
Oil	Capital-int. ag.	Nails	Capital-int. man.	Plums	Labour-int. ag.	Mercery	Unclassified
Ox Products	Capital-int. ag.	Oars	Capital-int. man.	Potatoes	Labour-int. ag.	Norwegian Goods	Unclassified
Oysters	Capital-int. ag.	Pipe Tools	Capital-int. man.	Resin	Labour-int. ag.	Provisions	Unclassified
Paint	Capital-int. ag.	Pipes	Capital-int. man.	Rice	Labour-int. ag.	Sailors	Unclassified
Pimento	Capital-int. ag.	Pistols	Capital-int. man.	Rubber	Labour-int. ag.	Ship Sully	Unclassified
Pollock	Capital-int. ag.	Pitch	Capital-int. man.	Rye	Labour-int. ag.	Soldiers	Unclassified
Quicksilver	Capital-int. ag.	Planks	Capital-int. man.	Safran	Labour-int. ag.	Swedish Goods	Unclassified
Raisins	Capital-int. ag.	Powder	Capital-int. man.	Spices	Labour-int. ag.	Waste	Unclassified
Rays	Capital-int. ag.	Sheet Metal	Capital-int. man.	Sugar	Labour-int. ag.		

Notes: List of 227 goods and classification into capital-intensive and labour-intensive agriculture and manufacturing, and unclassified goods.

A.A.2 Currencies and Units

Foreign currency is a non-issue in this data set as virtually all trade is recorded in Danish coins. In principle, the Soundtoll Data contain a variety of coins in which toll transactions are recorded. Over all recorded passages, there are 40 types of coins, of which 48.88% are Daler and 50.88% Skilling. I use data starting in 1668 that cover 4.38 million coin transactions. From 1668 onwards, 48.8% of coin transactions are recorded in Daler and 51.2% are in Skilling. Out of 4.38 million transactions, 347 use the Ort and one uses the Rosenobel, an English gold coin. I convert it based on its gold content. Thus, only 0.008% of transactions are in non-Danish currency. Table A.A.4 records conversion rates.

Table A.A.4: Currency conversion rates

	Daler	Skilling	Ort	Rosenobel
Daler	1	0.0104	0.25	3.94
Skilling		1	24	378
Ort			1	15.76
Rosenobel				1

Table A.A.5 lists the conversion of historical units of account into kilograms, liters, meters, and counts.

Table A.A.5: Conversion of units of account

Kilogram			Liter / Meter			Unit		
Unit	Commodity	Value	Unit	Commodity	Value	Unit	Commodity	Value
Boeter	Butter	3.96	Ahmer		150	Bale		1
Centner		50	Ancker		38.5	Ladning		1
Skippund		159	Boeter	Wine	585	Ries		0.1
Qtr. Skippund		39.75	Boeter	Oil	1225	Qtr. Bale		0.25
Drompt		35.6	Korff		24.8	Dbl. Bale		2
Faad		950	Drompt		431	Daler		1
Half Faad		475	Faad		930	Gulden		1
Qtr. Faad		237.5	Half Faad		465	Deger		10
Sacker		165	Qtr. Faad		232.5	Dosin		12
Half Sacker		82.5	Flaske		0.7245	Fierdinger		0.25
Tonder	Grain	100	Tonder	Beer	131	Kister	Glass	120
Tonder	Butter	112	Tonder	Salt	170	Kister	Tin	225
Half Tonder	Grain	50	Tonder	Herring	108	Half Kister	Glass	60
Half Tonder	Butter	56	Half Tonder	Beer	65.5	Half Kister	Tin	112.5
Kister	Tea	25	Half Tonder	Salt	85	Qtr. Kister	Glass	30
Half Kister	Tea	12.5	Half Tonder	Herring	54	Qtr. Kister	Tin	56.25
Qtr. Kister	Tea	6.25	Oxehoffde		232	Stocker		1
Pund		0.5	Oxehoffde	Beer Vin.	197	Half Stocker		0.5
Lispund		8	Moyer		876	Hundred		100
Last		1905	Ottinger		16.945	Ort		1
Last	Hemp, Hops	952	Piber		464	Paar		2
Skaeppe	Grain.	12.5	Qtr. Piber		116	Ring		240
Steen		15	Skaeppe		17.39	Skilling		1
Woger		17.856	Allen (m)		0.63	Schock		60
						Thilther		12
						Thommer		40
						Tylvt		12
						Thussend		1000

A.A.3 Duties

For every passage, the type and amount of duty is recorded, and usually there are several duties payable, some fixed and some proportional. As an example, I picked passage number 304. This ship paid 4 Daler of fire money ('fyrpenge', a fixed fee for the maintenance of the lighthouse) and passed the Sound on April 27, 1790. Its captain, Daniel Wendlandt, citizen of Windau, received a captain's compensation of 1.5 Daler, which was deducted from the total duty payable. This is a proportional tax free allowance. The first sub-total is 37.5 Daler, which I call proportional duty. When subtracting the captain's allowance, the second sub-total comes to 36 Daler, and the total duty payable is 40 Daler, where the fire money of 4 Daler has been added. The ship is

sailing with 75 last of rye; the recorded tax amount is 37.5 Daler, which equals the proportional duties and is the basis for my value calculations. The payable duties account for the allowance and the fixed fee. Table A.A.6 shows the summation exercise.

Table A.A.6: An example from the Soundtoll Data to show the different duties and sub-totals

Proportional duty	37.5 Daler
First sub-total	37.5 Daler
Captain's allowance	-1.5 Daler
Second sub-total	36 Daler
Fixed duty (fire money)	+4 Daler
Total payable duty	40 Daler

The Soundtoll team digitised the toll books in setting up three separate files.

1. The first one, called 'belastingen' or duties, records for each passage and line the type and amount of duty.

2. The second file, called 'ladingen' or carriage, records for each passage the origin and destination, the goods transported, and the overall amount of duty paid per good.

3. The third data set is called 'doorvaarten', or passages. For every passage, it records the day of passage through the Sound, information on the captain, discounts (or 'korting') applied to the duty, and two sub-totals and a total. A 'korting' is a discount, 87% of which are 'føring', a compensation for the captain in terms of a tax free allowance at their disposal.

The duty amount I use as a proxy for value is the proportional duty as in Waldinger (2022). Gøbel (2010) elaborates on the mechanics of the proportionality of this duty, which is summarised in Waldinger (2022) as a duty of 1-2% of cargo value. In order to discourage understating cargo value, a truth-telling mechanism was put in place: the toll official could purchase the cargo at the total value indicated in the customs forms. In Section A.A.4 below, I use the information on duty variation in Gøbel (2010) to back out underlying values from toll data by specifying a functional form for duty rates. This exercise produces virtually identical results, suggesting that it is appropriate to focus on the proportional duty.

A.A.4 Recovering underlying value

Underyling value, rather than the amount of duties paid, is a measure of value which I use as a robustness check. Gøbel (2010) notes that the rate at which proportional duties were applied varied by flag, with Danish ships paying lower tax rates, by good, with e.g. salt facing an additional duty, and also by time, with Swedish cargoes taxed at a lower rate only until 1720. Importantly, the origin and destination were not important for the determination of the proportional duty rate.

Motivated by this finding, I assume that taxed value, $\phi_{ijo,kt}$ is the product of a proportional duty $\psi_{o,kt}$, varying by flag o, sector k and time t, and underlying value $v_{ijo,kt}$:

$$\phi_{ijo,kt} = \psi_{o,kt} v_{ijo,kt} = \psi_{o,kt} p_{ijkt} q_{ijo,kt}, \tag{A.A.1}$$

such that the duty rate paid for shipping sector k goods from i to j under the flag of o in time t depend on the value of these goods multiplied by a proportional duty, $\psi_{o,kt}$, which varies only by flag o, sector k, and time t.

Turning equation A.A.1 into logs and absorbing log proportional duties by $d_{o,kt}$, I can recover proportional duties by estimating:

$$ln\phi_{ijo,kt} = d_{o,kt} + lnv_{ijo,kt}. \tag{A.A.2}$$

I let flag o be the country in which a ship is registered in the Soundtoll records.[3] Sectors k are the five sectors used throughout. While I estimate the following on annual data, I let the time dimension of fixed effects be decades. I set Denmark as the reference flag o, unclassified goods as the reference sector k, and the 1670s as the reference decade.[4]

Under the assumption that there are no time trends in the proportional duty rate of unclassified goods shipped under Danish flags, this recovers proportional duty rates by country, sector, and time. Choosing Denmark as the reference is motivated by the fact that Danish merchants paid the lowest rates throughout. With these estimated tariff rates, I turn the observed tariff amounts into underlying value by dividing toll amounts by the normalised duty rate, in line with equation A.A.1.

[3]The main countries are Sweden, Russia, the German states, Denmark, the Netherlands, France, and Britain. Within these larger states, I also allow Norway, Belgium, Estonia, Lithuania, Latvia and Finland to potentially have different duty rates.

[4]Using $d_{o,kt} = ln\psi_{o,kt}$ and writing in terms of $\psi_{o,kt}$, estimating equation A.A.2 recovers:

$$\frac{\psi_{ikt}\psi_{i'kt'}\psi_{ik't'}\psi_{i'k't}}{\psi_{ikt'}\psi_{i'kt}\psi_{ik't}\psi_{i'k't'}}.$$

Appendix A.B History and Geography Appendix

Only ships passing the Oresund appear in the data. Thus, trade flows between cities on either side of the Sound (say, between London and Amsterdam) is unobserved. For these flows, I do not impose zero trade, but rather I will not include these. The Soundtoll data, however, cover almost all trade between the Baltic and North Sea areas.

It is highly unlikely that any trade between, say, Danzig and London occurred other than through the Sound. Other straits, such as the Little Belt between Jutland and Funen, were difficult and dangerous to navigate for larger ships such as the ones used for trading with the West. The main alternative to passing the Sound, between Sweden and Denmark's main island Zealand, was sailing through the Great Belt, between Zealand and Funen. Degn (2017) makes a number of arguments against this being an issue for these data. The Danish king introduced a prohibition on sailing through the Belt, directed specifically against the Prussian towns, extending this prohibition later on to all foreign nations. This was precisely because the Great Belt potentially allowed merchants to circumvent the Sound Toll. Only expert skippers possessed the specific knowledge required to pass through the Great Belt, a route usually taken in case of storms. Degn (2017) concludes that skippers only reluctantly avoided the Sound. Finally, toll records from the Great Belt reveal that the number of passages was small. There is also no reason to believe these were high trade value passages: according to Gøbel (2010), the toll revenue from the Great Belt and the Little Belt combined equalled but a few percent of the Sound Toll revenues. Alternatively, trade over land faced a high number of tolls, too, in that many more borders would have to be crossed when compared to passing through the Sound, compounding already high transport costs over land.

My source for identifying cities in the trade data are two files provided by the Soundtoll team, which convert 90,737 original city names to 3,085 unified city names. My sample are coastal cities (less than 25 kilometres distance to the coast, or those along a major river up to 100 kilometres from the coast) between Ireland to the West, Saint Petersburg to the East, the Arctic Circle to the North, and Bayonne to the South. 1,425 cities are in this sample, though only 676 are actively trading during the period I study.

A.B.1 Classification of Areas

The trade regressions in Section 1.3 include area-time fixed effects. These areas are not just one level of European NUTS regions as their sizes vary greatly and as Russia and Norway are not EU members. NUTS 1 is too aggregate for Northern Europe, with Finland and Denmark as

each only one region. NUTS 2 is sufficiently granular in Northern Europe, but too granular in Western Europe, prompting a sixth of areas to only contain one port city. At NUTS 3, a third of all areas contain only one port.

Therefore, I construct similarly-sized areas: I use NUTS 2 (or a close equivalent) throughout, other than for the United Kingdom, France, Germany, and Spain, where I use NUTS 1. When aggregating areas below, this is always to ensure that no area has only one port city. Table A.B.1 provides an overview of the number of units and their NUTS correspondance by country. Figure A.C.9 shows that the export expansion result goes through also when including NUTS 1-time and NUTS 2-time fixed effects instead of the area-time fixed effects proposed here.

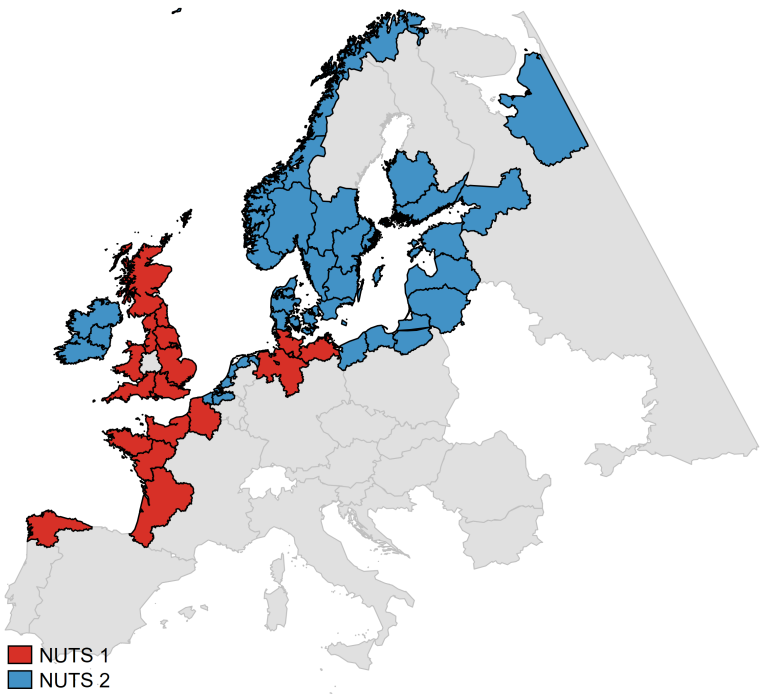
Table A.B.1: Areas and NUTS correspondence

Country	Unit type	# (in data)	NUTS level
Belgium	Provinces	2	NUTS 2
Denmark	Regions	5	NUTS 2
England	Regions (pooled)	7	NUTS 1
Estonia	Country	1	NUTS 2
Finland	Provinces (1997)	2	∼NUTS 2
France	Regions	5	NUTS 1
Germany	States	5	NUTS 1
Ireland	Provinces (incl. NI)	3	∼NUTS2
Latvia	Country	1	NUTS 2
Lithuania	Country	1	NUTS 2
Netherlands	Province	5	NUTS 2
Norway	Traditional regions	5	∼NUTS 2
Poland	Voivodeships	3	NUTS 2
Russia	Oblasts	2	∼NUTS 2
Scotland	Country	1	NUTS 1
Spain	Pooled	1	NUTS 1
Sweden	National areas (pooled)	6	NUTS 2
Wales	Country	1	NUTS 1

The United Kingdom is disaggregated into Scotland and Wales, and England is disaggregated into her nine regions, where I pool the East Midlands with the East of England, London with the South East, and the Isle of Man with the North West, and add the Channel Islands as an area. Ireland's four provinces are aggregated to three together with Northern Ireland, with Connacht pooled with Ulster. The Netherlands are made up of their 5 provinces by the North Sea. Belgium's areas are her three provinces by the North Sea, where Antwerpen is grouped with East Flanders. Denmark has 5 regions and the Faroe Islands are pooled with the main country. For Finland, I pick the provinces of Finland as of 1997, pooling Oulu and Lapland as Northern Finland and Aland with Western Finland. Only Southern and Western Finland show up in the data, however. France's areas are the five regions on the Atlantic Coast and Channel.

For Germany, I pick the current 16 states, of which 5 show up in the data. The three Baltic states are separate areas. Norway is disaggregated into five traditional regions that show up in the data. For Russia, I choose the two Baltic oblasts, pooling Arkhangelsk with Leningradskaya Oblast. Eight national areas are chosen in Sweden, corresponding to the NUTS 2 areas, though only 6 show up in the data. The few ports in Northern Spain, specifically in Asturias and Galicia, are pooled as one area. Finally, the areas in modern-day Poland are the three voivodeships on the Baltic Sea.

Figure A.B.1: Areas



Notes: This map shows the 56 areas into which regions are aggregated. For the UK, Germany, France, and Spain, I choose NUTS 1 regions, and for the rest of the sample areas that are equivalent to NUTS 2 regions. This avoids areas with an individual port city and generates areas of comparable sizes.

A.B.2 Additional Plague Results

Figure A.B.2 shows digitised army marching routes. As armies spread the plague across Europe, regions' proximity to these routes is used in Table A.B.4 to predict mortality rates when this information is not known. Table A.B.2 shows information on plague outbreaks in my sample and provides an overview of timing and mortality rates. For plagued cities whose mortality rates were not observed, I form predicted mortality rates from regressing geographical covariates, the timing of the plague, and the proximity of army marching routes on observed mortality routes. These results are shown in Table A.B.4 and used to form predicted mortality rates. Army marching are digitised based on maps by Spruner & Menke (1880) and Barraclough (1997). Table A.B.4

regresses geographical controls, the timing of the plague, and proximity to army marching routes on observed mortality rates for the subset of plagued cities. For half of these cities, I form predicted mortality rates based on the results presented in this table, as no mortality estimates are available.

Figure A.B.2: Army marching routes, 1706-1714



Notes: Map of army marching routes between 1706 and 1714, when the plague was recorded to have spread across Eastern and later Northern and Central Europe. No distinction is made between armies, as both allied and Russian troops are known to have spread the plague. I base these routes on maps by Spruner & Menke (1880) and Barraclough (1997) which I digitised.

Table A.B.2: List of plagued cities and mortality estimates

City	Modern Country	Plague	Time	Population	Mortality	Source
København/Copenhagen	Denmark	1	1711	60000	.283	Kroll & Grabinsky (2007), Kroll (2006)
Helsingør	Denmark	1	1710 - 1711	4000	.408	Frandsen (2009)
Flensburg	Germany	1	not specified			Ulbricht (2004)
Frederiksort	Germany	1	1712			Ulbricht (2004)
Kiel	Germany	1	1712			Frandsen (2009)
Rendsburg	Germany	1	1712			Ulbricht (2004)
Altona	Germany	1	1712			Frandsen (2009), Winkle (1983)
Glückstadt	Germany	1	1712			Frandsen (2009)
Pinneberg	Germany	1	1712			Winkle (1983)
Itzehoe	Germany	1	1712			Frandsen (2009)
Schleswig	Germany	1	1712			Ulbricht (2004)
Helsinki	Finland	1	1710		.66	Engström (1994)
Rauma	Finland	1	1710 - 1711			Vourinen (2007)
Turku	Finland	1	1710 - 1711	6000	.33	Vourinen (2007)
Raseborg	Finland	1	1710			Kroll & Grabinsky (2007), Kroll (2006)
Pori	Finland	1	1710 - 1711			Vourinen (2007)
Oulu	Finland	1	1710 - 1711			Vourinen (2007)
Borgå	Finland	1	1710			Vourinen (2007)
Pietarsaari	Finland	1	1710 - 1711			Vourinen (2007)
Kokkola	Finland	1	1710 - 1711			Vourinen (2007)
Uusikaupunki	Finland	1	1710 - 1711			Vourinen (2007)
Enköping	Sweden	1	1710			Frandsen (2009)
Stockholm	Sweden	1	1710 - 1711	53750	.378	Kroll & Grabinsky (2007), Kroll (2006)
Visby	Sweden	1	1710 - 1711	2375	.233	Kroll & Grabinsky (2007), Kroll (2006)
Karlskrona	Sweden	1	1710 - 1712			Persson (2011)
Karlshamn	Sweden	1	1710 - 1712			Persson (2011)
Jönköping	Sweden	1	1710 - 1711	2500	.374	Kroll & Grabinsky (2007), Kroll (2006)
Ystad	Sweden	1	1712	1950	.385	Kroll & Grabinsky (2007), Kroll (2006)
Blekinge	Sweden	1	1710			Persson (2011)
Malmö	Sweden	1	1712	5000	.35	Kroll & Grabinsky (2007), Kroll (2006)
Linköping	Sweden	1	1710 - 1711	1500	.295	Kroll & Grabinsky (2007), Kroll (2006)
Domsten	Sweden	1	1711			Kroll & Grabinsky (2007), Kroll (2006)
Narva	Estonia	1	1710 - 1711	3000		Kroll & Grabinsky (2007), Kroll (2006)
Tallinn/Reval	Estonia	1	1710	9900	.661	Kroll & Grabinsky (2007), Kroll (2006)
Kuressaare	Estonia	1	1710			Frandsen (2009)
Saaremaa	Estonia	1	1710			Frandsen (2009)
Riga	Latvia	1	1710 - 1711	10477	.651	Kroll & Grabinsky (2007), Kroll (2006)
Pärnu	Estonia	1	1710	2350	.529	Kroll & Grabinsky (2007), Kroll (2006)
Kaliningrad/Königsberg	Russia	1	1709 - 1710	36250	.239	Kroll & Grabinsky (2007), Kroll (2006)
Klaipeda/Memel	Lithuania	1	1709 - 1710			Kroll & Grabinsky (2007), Kroll (2006)
Sovjetsk/Tilsit	Russia	1	1709 - 1710			Kroll & Grabinsky (2007), Kroll (2006)
Gdansk/Danzig	Poland	1	1709	50000	.533	Kroll & Grabinsky (2007), Kroll (2006)
Elblag/Elbing	Poland	1	1709 - 1710		.3	Frandsen (2009)
Kamien Pomorski/Cammin	Poland	1	not specified			bei der Wieden (1999)
Stargard	Poland	1	1710 - 1711	7000	.041	Kroll & Grabinsky (2007), Kroll (2006)
Szczecin/Stettin	Poland	1	1709 - 1711	11250	.171	Kroll & Grabinsky (2007), Kroll (2006)
Wolin	Poland	1	1710 - 1711			bei der Wieden (1999)
Anklam	Germany	1	not specified			bei der Wieden (1999)
Greifswald	Germany	1	1711			bei der Wieden (1999)
Wolgast	Germany	1	1710 - 1711		.4	bei der Wieden (1999)
Stralsund	Germany	1	1710 - 1711	7250	.314	Kroll & Grabinsky (2007), Kroll (2006)
Goleniow/Gollnow	Poland	1	1709			von Schöning (1837)
Hamburg	Germany	1	1712 - 1714	70000	.130	Kroll & Grabinsky (2007), Kroll (2006)
Stade	Germany	1	1712			Frandsen (2009), Winkle (1983)
Bremen	Germany	1	1712 - 1713	28000	.007	Frandsen (2009)

Notes: Sources are indicated in the last column. Population and mortality estimates, in the case of several available sources, denote the mean estimate. The modern names and countries of cities are used here.

Table A.B.3: List of besieged cities

City	Modern Country	Siege
København	Denmark	1
Tönning	Germany	1
Hamburg	Germany	1
Halden	Norway	1
Gävle	Sweden	1
Härnösand	Sweden	1
Hudiksvall	Sweden	1
Piteå	Sweden	1
Söderhamn	Sweden	1
Umeå	Sweden	1
Södertälje	Sweden	1
Norrköping	Sweden	1
Nyköping	Sweden	1
Trosa	Sweden	1
Sundsvall	Sweden	1
Vyborg	Russia	1
Narva	Estonia	1
Riga	Latvia	1
Szczecin	Poland	1
Wolgast	Germany	1
Stralsund	Germany	1
Wismar	Germany	1

Notes: Six cities are both besieged and plagued. The six cities are Copenhagen (København), Narva, Riga, Stettin (Szczecin), Stralsund, and Wolgast. For Copenhagen and the first siege of Riga, there was no association with the plague, as these sieges occurred at the very beginning of the war in 1700 before the plague spread.

Table A.B.4: Determinants of urban mortality

	Mortality rate
	(1)
Year of plague outbreak	-0.000
	(0.001)
Latitude	0.010
	(0.022)
Longitude	0.017**
	(0.007)
Distance to closest army	-0.062**
	(0.023)
East of the Sound	0.160**
	(0.055)
Observations	19

Notes: Standard errors in parentheses. * p <0.10, ** p <0.05, *** p <0.01. The dependent variable is the median urban mortality estimate. The independent variables are the year of the plague outbreak, latitude, longitude, distance to the closest army route between 1706 and 1714, and a dummy for the city being to the East of the Sound. The rationale for including army marching routes is that armies spread the plague.

A.B.3 Demography

Table A.B.5 shows that plagued cities were not larger or smaller than other cities, taking into account time trends and locational fundamentals by including time (or country x time) and city fixed effects. The source for these city-level populations is Buringh (2021). As the closest data point after the plague is in 1750, I conclude that cities' populations had recovered to the trend after four decades. For regional rather than urban populations, I draw on Swedish parish-level population data (Andersson Palm (2025)). In Table A.B.6, I find that by 1750, the populations of plagued Swedish parishes had recovered.

Table A.B.5: Impact of plague on city population

	Population		Log Population	
	(1)	(2)	(3)	(4)
Plague x Year=800	0.0749	-0.0242	0.224	0.154
	(11.86)	(10.77)	(0.811)	(0.584)
Plague x Year=900	0.0715	-0.0998	0.141	0.0493
	(11.86)	(10.77)	(0.810)	(0.625)
Plague x Year=1000	0.251	-0.0178	0.0817	0.0286
	(11.86)	(10.76)	(0.737)	(0.531)
Plague x Year=1100	0.161	-0.232	0.112	0.0373
	(11.84)	(10.74)	(0.710)	(0.479)
Plague x Year=1200	0.424	-0.163	0.130	0.0898
	(11.82)	(10.72)	(0.697)	(0.469)
Plague x Year=1300	1.339	0.352	0.366	0.282
	(11.78)	(10.68)	(0.691)	(0.460)
Plague x Year=1400	1.772	1.043	0.460	0.468
	(11.74)	(10.64)	(0.690)	(0.458)
Plague x Year=1500	3.816	2.709	0.597	0.592
	(11.35)	(10.34)	(0.688)	(0.455)
Plague x Year=1550	4.758	3.489	0.622	0.637
	(11.24)	(10.23)	(0.686)	(0.454)
Plague x Year=1600	6.464	4.798	0.684	0.704
	(11.05)	(10.02)	(0.686)	(0.455)
Plague x Year=1650	7.789	5.332	0.783	0.747*
	(10.81)	(9.784)	(0.685)	(0.453)
Plague x Year=1700	9.645	6.381	0.800	0.737
	(10.52)	(9.499)	(0.687)	(0.455)
Plague x Year=1750	11.20	7.018	0.840	0.668
	(10.29)	(9.271)	(0.686)	(0.454)
Plague x Year=1800	14.78	8.766	0.954	0.697
	(9.583)	(8.657)	(0.687)	(0.454)
Plague x Year=1850	21.37**	10.64	0.844	0.692
	(8.999)	(8.147)	(0.687)	(0.454)
Plague x Year=1900	63.51***	44.93**	0.767	0.948**
	(19.85)	(18.64)	(0.691)	(0.460)
Plague x Year=1950	107.4**	87.99**	0.548	0.907*
	(42.88)	(38.62)	(0.695)	(0.465)
Plague x Year=2000	162.9***	153.2***	0.438	0.999**
	(59.85)	(54.14)	(0.694)	(0.467)
Constant	19.54***	19.69***	1.640***	1.637***
	(0.940)	(0.947)	(0.0266)	(0.0181)
Fixed Effects:				
– City	✓		✓	
– Country x Year	✓	✓	✓	✓
– Year		✓		✓
Observations	22,470	22,489	17,027	17,059

Notes: Robust standard errors in parentheses. * p <0.10, ** p <0.05, *** p <0.01. The dependent variable is (log) city population. The independent variable is a plague dummy interacted with year dummies. The area is restricted to the Baltic Sea and North Sea areas, specifically to cities between the tip of Norway, the Northern end of the Alps, the Irish West Coast and Arkhangelsk.

Table A.B.6: Parish-level population recovery after the plague

	Population	Log Population
	(1)	(2)
Plague & Year=1750	25.51	0.0337***
	(28.84)	(0.0125)
Plague & Year= 1810	-28.01	-0.0324**
	(59.52)	(0.0151)
Fixed Effects:		
– Parish	✓	✓
– Year	✓	✓
Observations	7,032	7,032

Notes: Standard errors clustered at the parish level in parentheses. * p <0.10, ** p <0.05, *** p <0.01. The dependent variables are parish-level (logarithmic) populations in 1690, 1750, and 1810. The independent variable is a post-plague dummy.

Demographic Response

What was the demographic response to this plague outbreak? Theoretically, three main channels could contribute to population recovery. First, an increase in fertility. Guinnane (2011) argues for a positive elasticity of fertility with respect to income. However, there is reason to doubt the strength and direction of this relationship in the particular area studied here. There is evidence for delayed marriage (De Moor & van Zanden (2009)) after the Black Death for the North Sea region, which may suggest falling fertility. While there is no direct evidence on Baltic cities, this area falls under the European Marriage Pattern (Hajnal (1965)).

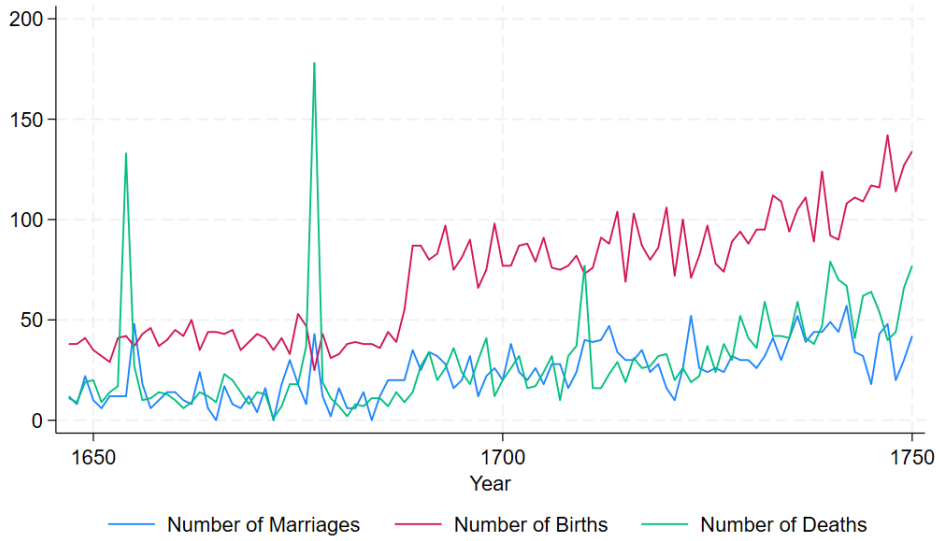
Population growth may still have accelerated due to lowered mortality, the second channel. Waldinger (2022) shows that higher agricultural productivity was associated with lower mortality. Migration is the third channel through which urban populations may have recovered. Waldinger (2022) also shows for England that warming increased agricultural productivity and drew in more migrants. While there is limited evidence for such migration across regions, I argue that rural-to-urban migration within regions may be sufficient to produce urban population recovery.

The Scanian Economic Demographic Database provides a uniquely granular source to answer this question (Bengtsson et al. (2014)). Figure A.B.3 shows annual births, deaths, and marriages for all of Scania. The 1710 peak in deaths is the plague outbreak studied in this paper. It was followed by a slight increase in marriages and births. This suggests a positive fertility response.

I also present results on the fertility channel at the parish level in Table A.B.7. I find no significant fertility response at the parish level. As about a third of inhabitants of these parishes vanished, this implies a positive fertility response. However, this result is based on 59 parishes in

one particular region of Sweden, limiting the external validity of this finding.

Figure A.B.3: Number of births, deaths, and marriages in Scania, 1647-1750



Notes: The 1650s peak in deaths marks the Second Northern War and the 1670s peak the Scanian War.

Table A.B.7: Impact of plague on number of births in 59 Scania parishes

	Births		Log Births	
	(1)	(2)	(3)	(4)
Post Plague	0.174	-0.376	-0.065	-0.104
	(0.725)	(0.735)	(0.068)	(0.083)
Fixed Effects:				
– Origin	✓	✓	✓	✓
– Year	✓	✓	✓	✓
Years	All	1647-1750	All	1647-1750
Observations	1,015	591	973	578

Notes: Standard errors in parentheses. * p <0.10, ** p <0.05, *** p <0.01. The dependent variable is a the (log) number of births per year and parish. The independent variable is a plague dummy, equal to one after the parish suffered a plague outbreak.

Insights from Other Plague Outbreaks

The Great Northern War plague outbreak's mortality rate of 36% is similar to the 40% mortality rate of the Black Death (Jedwab et al. (2022)). A key difference is that the Great Northern

War plague outbreak was geographically very concentrated. This is a feature of later plague outbreaks also in Southern Europe (Alfani (2023)). Several reasons may have contributed to this, most importantly environmental and epidemiological changes affecting both immunity and transmissibility (Alfani & Murphy (2017)). It is therefore not atypical for such a late plague outbreak to be geographically concentrated.

The economic consequences of plague outbreaks depend crucially on how mortality interacts with sex, age, and socioeconomic status. Alfani & Murphy (2017) summarises the existing literature on other plague outbreaks. There is no evidence that mortality differed by sex and ambiguous evidence on age. For the original Black Death and some later plague outbreaks, there is clear evidence that elites and lower-class citizens alike suffered high mortality.[5] Raster (2023) analyses plague mortality for the Great Northern War plague outbreak specifically. While geographically constrained to Estonia, he finds no significant differences in mortality rates. Frandsen (2009) shows qualitative evidence confirming this finding. Thus, the compositional effects of the plague outbreak should have been very limited.

[5]Pullan (1992) describes that 17% of the members of the Great Council were killed in the 1630 Venice plague and that 40% of the members of the Great and Low Councils died in the 1656-57 Genova plague outbreak.

Appendix A.C Trade Appendix

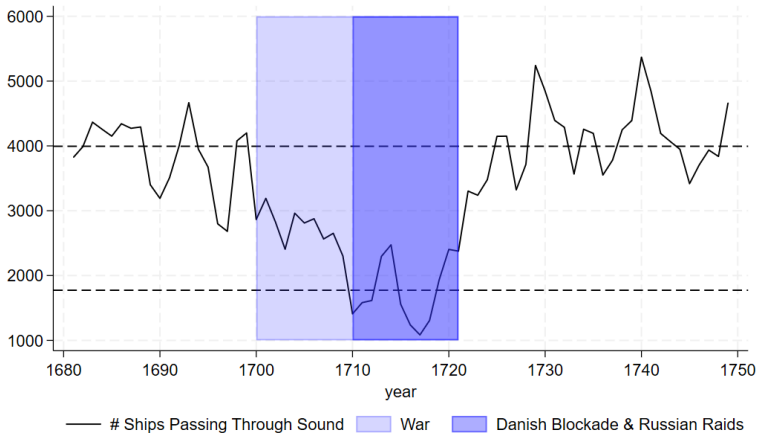
This Appendix holds all additional and robustness results for trade. The structure of this Appendix follows the structure of the trade results in the main paper. Figure A.C.1 shows how blockades and naval raids affected the number of ships passing through the Sound. Table A.C.1 lists blocked and raided Baltic ports, which is used in Table A.C.2 to show that proximity to the naval base from which the Danish blockade was enforced predicts significantly lower exports, beyond blockades and raids.

Table A.C.1: Naval raids and blockades affecting Baltic ports, 1709-1721

City	Event	Year(s)	Source
Christiansø	Danish base for blockade	1709–1720	Andersen (2004)
Karlskrona	Allied blockade	1716–1720	Aldridge (2009)
Gothenburg	Danish blockade	1717–1719	Andersen (2004)
Marstrand	Danish blockade	1719	Andersen (2004)
Strömstad	Danish blockade	1719	Andersen (2004)
Öregrund	Russian coastal raid	1719	Ericson Wolke (2012)
Gotland	Russian coastal raid	1717	Ericson Wolke (2012)
Norrtälje	Russian coastal raid	1719	Ericson Wolke (2012)
Södertälje	Russian coastal raid	1719	Ericson Wolke (2012)
Stockholm archipelago	Russian coastal raid	1719	Ericson Wolke (2012)
Trosa	Russian coastal raid	1719	Ericson Wolke (2012)
Nyköping	Russian coastal raid	1719	Ericson Wolke (2012)
Norrköping	Russian coastal raid	1719	Ericson Wolke (2012)
Gävle	Russian coastal raid	1719	Ericson Wolke (2012)
Umeå	Russian coastal raid	1720	Ericson Wolke (2012)
Söderhamn	Russian coastal raid	1721	Ericson Wolke (2012)
Hudiksvall	Russian coastal raid	1721	Ericson Wolke (2012)
Sundsvall	Russian coastal raid	1721	Ericson Wolke (2012)
Härnösand	Russian coastal raid	1721	Ericson Wolke (2012)
Piteå	Russian coastal raid	1721	Ericson Wolke (2012)

Notes: Based on the specifics sources in column 4 and on Christiansø Administration (2020), Ullman (2006), Frost (2000), and Kuylenstierna (1899).

Figure A.C.1: Number of ships passing the Sound per year



Notes: This Figures displays the number of ships passing the Sound each year. In light blue, the war years of 1700 to 1721 are shaded; in dark blue, the years of 1710 to 1721 are shaded, during which the Danish navy blocked ships between Sweden and its Northern German posessions and the Russian army raided Baltic coastal areas and frequently intercepted ships. During this second half of the conflict, the number of ships passing the Sound collapsed to about a third, reducing observed trade for both plagued and unplagued ports.

Table A.C.2: Proximity to Christiansø is associated with significantly lower exports

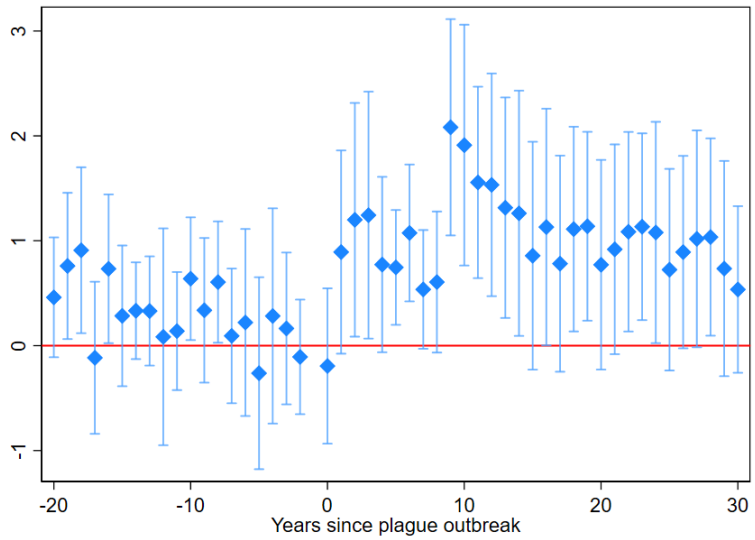
	Log Proximity to Christiansø			
	(1)	(2)	(3)	(4)
Logarithmic Proximity to Christiansø	-0.425**	-0.581***	-0.455***	-0.604***
	(0.166)	(0.210)	(0.172)	(0.217)
Logarithmic Proximity x Plague		0.022		0.021
		(0.014)		(0.013)
Dummy for Danish Blockade			-0.160	-0.112
			(0.137)	(0.130)
Dummy for Russian Raid			-0.443	-0.429
			(0.316)	(0.319)
Fixed Effects:				
– Origin	✓	✓	✓	✓
– Area x Year	✓	✓	✓	✓
Observations	13,416	13,416	13,416	13,416

Notes: Standard errors clustered at the origin level in parentheses. * p <0.10, ** p <0.05, *** p <0.01. The dependent variable is the number of ships leaving an origin port. The independent variables are logarithmic proximity to the naval base from which the Danish blockade of Swedish ports was operated, defined as the product of a war period indicator (equal to one for 1709–1721) and the negative log of the great-circle distance in kilometres from each city to Christiansø, so that it takes larger values for cities that are geographically closer to the Danish blockade base during the war years and is zero in peacetime. Further, the interaction of this proximity measure with a plague dummy, a blockade dummy for additionally directly blocked ports, and a dummy for ports raided by Russian troops. All three are based on sources listed in Table A.C.1.

A.C.1 Fact #1: Capital-intensive exports increase more than labour-intensive exports.

Figure A.C.2 shows that reassigning four debatable goods (wine, tow, ash, and planks) as labour-intensive produces virtually identical results. Table A.C.3 shows PPML and OLS results for Fact #1 in a difference-in-difference-in-difference set up. Table A.C.4 repeats the specification while including army proximities and sieges. Figure A.C.3 shows findings by factor intensity using the imputation estimator in Borusyak, Jaravel & Spiess (2021). Table A.C.5 corroborates the increase in capital-intensive exports over labour-intensive exports by sector. Figure A.C.4 shows the result by goods. Figure A.C.5 shows that there is no shift into land-intensive exports after the plague. Finally, Figure A.C.6 shows that the shift into capital-intensive exports lasts for almost 90 years.

Figure A.C.2: Alternative assignment as capital-intensive: Fact #1



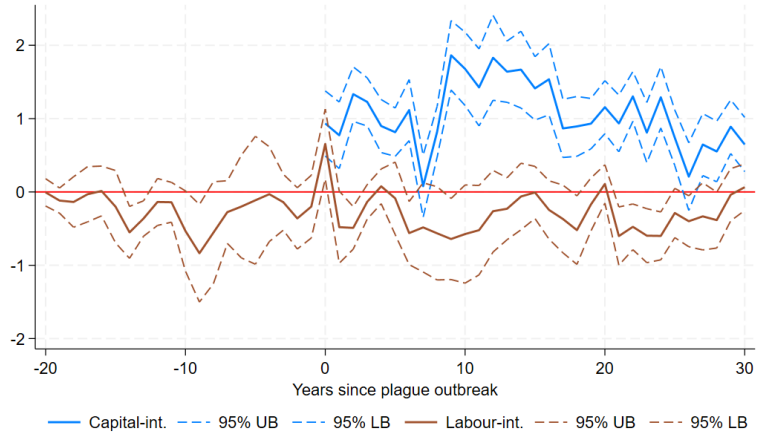
Notes: Among the top 30 goods (see Table A.A.2), I have reassigned four debatable goods (wine and tow from CA to LA, planks and ash from CM to LM) and find almost identical results.

Table A.C.3: Plague increases capital-intensive exports more than labour-intensive exports

	PPML		OLS	
	(1)	(2)	(3)	(4)
Post Plague	-1.219**	-1.390***	-0.570***	-0.578***
	(0.484)	(0.436)	(0.205)	(0.186)
Post Plague x Capital-Int.	0.904*	1.385***	0.553*	0.929***
	(0.527)	(0.229)	(0.289)	(0.164)
Fixed Effects:				
– Origin x Sector	✓	✓	✓	✓
– Sector x Year	✓		✓	
– Year		✓		✓
Estimator	PPML	PPML	OLS	OLS
Observations	68,989	71,980	12,886	12,901

Notes: Difference-in-differences set up dropping the war years of 1700 to 1721. Standard errors clustered at the origin level in parentheses. * p <0.10, ** p <0.05, *** p <0.01. The dependent variable is annual exports by sector in levels (columns 1-2) and logs (columns 3-4). The independent variable is a post plague dummy, equal to one after the origin suffered a plague outbreak, and a dummy interacting the post plague dummy with a dummy for capital-intensive exports. Columns 1-2 show unweighted results, whereas in columns 3-4 the weight are exports in levels. Annual growing season temperatures are an additional control in all specifications.

Figure A.C.3: Fact #1 using imputation estimator in Borusyak, Jaravel & Spiess (2021)



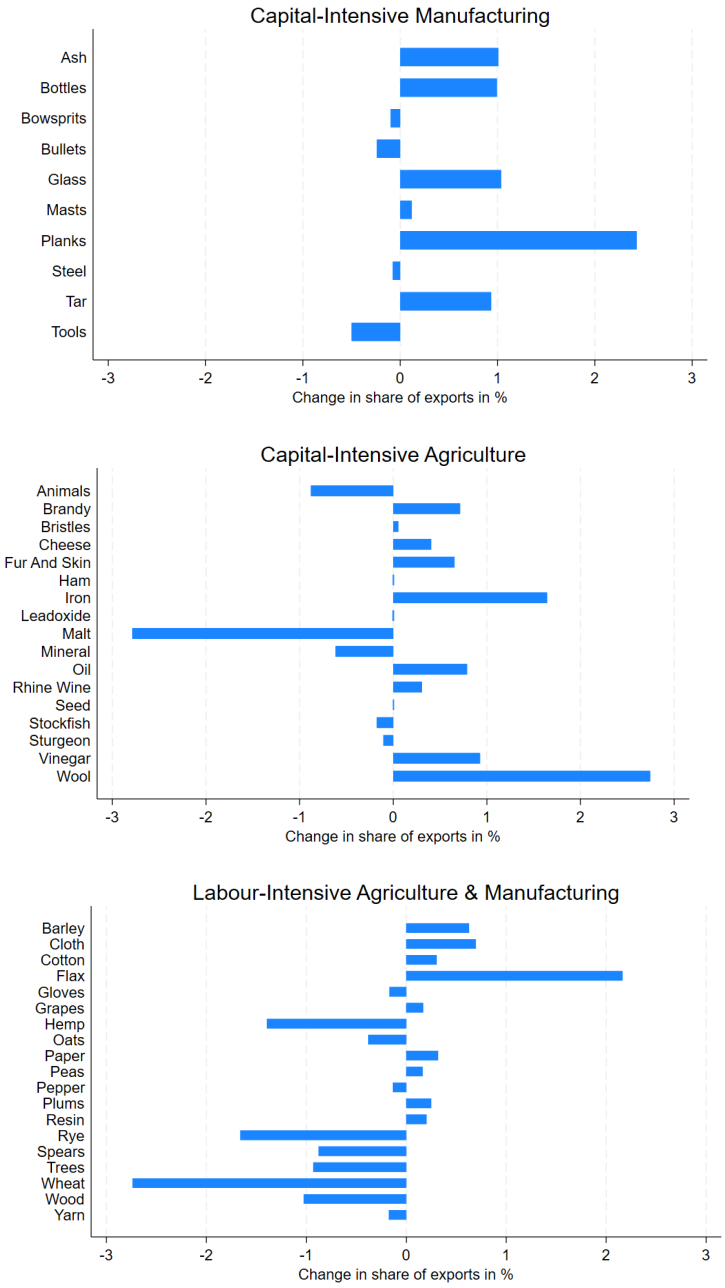
Notes: I estimate equation 1.3.1 by factor intensity using the imputation estimator in Borusyak, Jaravel & Spiess (2021). The package does not permit the estimation of pretrend heterogeneity, so these estimates are pooled across both factor intensities.

Table A.C.4: Shift into capital-intensive exports by army proximity and siege

	PPML		OLS		PPML		OLS	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Army Distance x Post War	-0.060	-0.061	0.015	0.006				
	(0.071)	(0.057)	(0.031)	(0.025)				
Army Distance x Post War x Capital-Int.	0.061	0.047**	0.008	-0.002				
	(0.043)	(0.023)	(0.028)	(0.016)				
Siege x Post War					0.715***	0.750***	0.579***	0.607***
					(0.098)	(0.083)	(0.090)	(0.092)
Siege x Post War x Capital-Int.					0.335	0.309	0.007	0.007
					(0.273)	(0.315)	(0.229)	(0.239)
Post Plague	-1.136**	-1.256***	-0.424**	-0.545***	-1.156***	-1.321***	-0.615***	-0.709***
	(0.510)	(0.419)	(0.177)	(0.119)	(0.385)	(0.324)	(0.131)	(0.103)
Post Plague x Capital-Int.	1.132**	1.265***	0.514*	0.842***	1.109**	1.345***	0.573*	0.923***
	(0.530)	(0.270)	(0.308)	(0.193)	(0.563)	(0.340)	(0.329)	(0.216)
Fixed Effects:								
– Origin x Sector	✓	✓	✓	✓	✓	✓	✓	✓
– Sector x Year	✓		✓		✓		✓	
– Year		✓		✓		✓		✓
Estimator	PPML	PPML	OLS	OLS	PPML	PPML	OLS	OLS
Observations	100,530	103,750	17,353	17,368	102,636	105,908	17,545	17,560

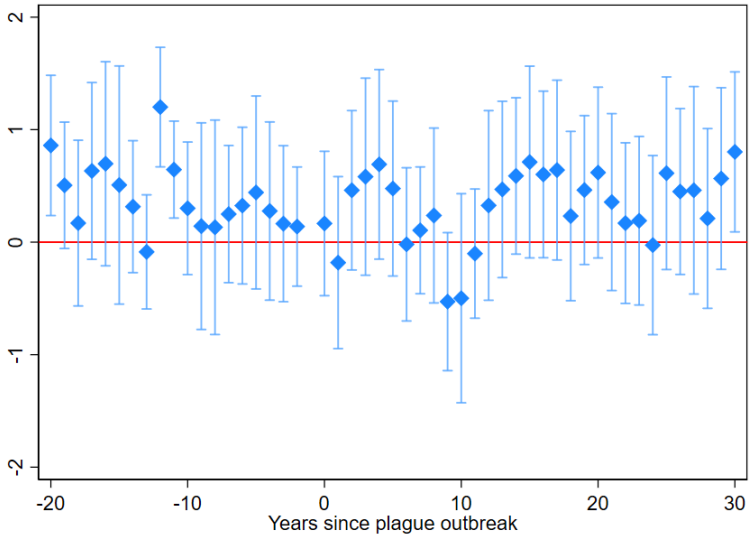
Notes: Standard errors clustered at the origin level in parentheses. * p <0.10, ** p <0.05, *** p <0.01. The dependent variable is annual exports by sector in levels (PPML) or logs (OLS). The independent variable is a post plague dummy, equal to one after the origin suffered a plague outbreak, and a dummy interacting the post plague dummy with a dummy for capital-intensive exports. Additionally, in columns 1-4 I include the smallest distance any army had to a region during the Great Northern War, interacted with a post-war dummy (equal to 1 after 1721), and this interactions interaction with a capital-intensity dummy. The PPML columns show unweighted results, whereas in the OLS columns the weight are exports in levels. Annual growing season temperatures are an additional control in all specifications. Table A.B.3 shows besieged cities and Figure A.B.2 shows army routes.

Figure A.C.4: Fact #1: Shift into capital-intensive exports by good



Notes: Results from regressing a plague dummy on the share a good constitutes among origin-time exports, including origin and time fixed effects. Shown is a selection of point estimates that are significant at the 10% level.

Figure A.C.5: No shift into land-intensive exports after the plague



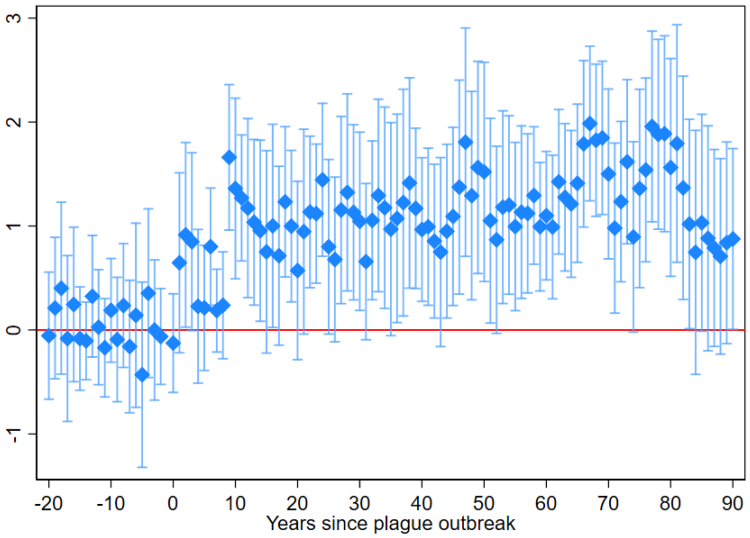
Notes: Distinguishing land-intensive (arable crops, livestock and animal products, plantation goods, raw forestry) from labour-intensive goods (crafted manufactures, iron and metalwork, textiles and clothing).

Table A.C.5: Plague increases capital-intensive exports more than labour-intensive exports in both sectors

	Manufacturing		Agriculture	
	(1)	(2)	(3)	(4)
Post Plague	-1.491***	-0.547***	-1.136***	-0.566***
	(0.498)	(0.121)	(0.355)	(0.183)
Post Plague x Capital-Int.	1.267***	0.787***	1.507**	0.740*
	(0.235)	(0.142)	(0.647)	(0.379)
Fixed Effects:				
– Origin x Sector	✓	✓	✓	✓
– Year	✓	✓	✓	✓
Estimator	PPML	OLS	PPML	OLS
Observations	31,023	5,043	53,369	9,578

Notes: Standard errors clustered at the origin level in parentheses. * p <0.10, ** p <0.05, *** p <0.01. The dependent variable is annual exports by sector in levels (columns 2 and 4) and logs (columns 1 and 3). The independent variable is a post plague dummy, equal to one after the origin suffered a plague outbreak, and a dummy interacting the post plague dummy with a dummy for capital-intensive exports. Columns 1 and 3 show unweighted results, whereas in columns 2 and 4 the weight are exports in levels. Annual growing season temperatures are an additional control in all specifications.

Figure A.C.6: Capital-intensive exports expand relative to labour-intensive exports also in the long run

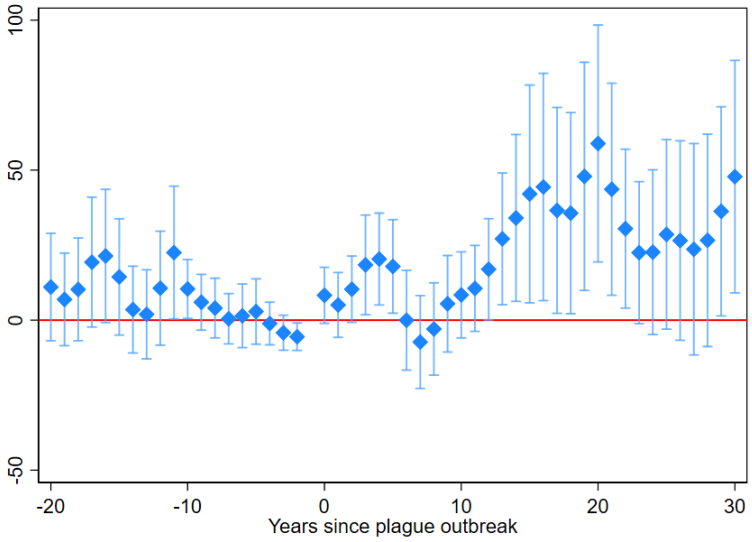


Notes: Estimation of equation 1.3.1 in the long run. Standard errors clustered at the origin level.

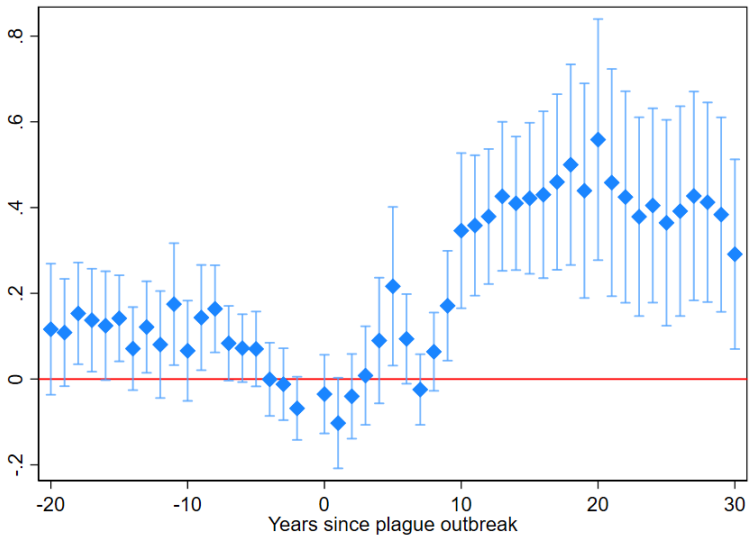
A.C.2 Fact #2: Plagued regions capture larger shares of destination markets.

I now present robustness results for the post-plague export expansion. Figure A.C.7a shows that export volumes expand and Figure A.C.7b shows the extensive margin expansion. Table A.C.6 shows results from a difference-in-differences set up excluding the war years of 1700 to 1721. Figure A.C.8a extends the pre-period to 30 years and again finds no pre-trends. Figure A.C.11a adds lagged temperatures. I also show results from cleaned values following the decomposition of duty records into underlying value as outlined in Section A.A.4 in Figure A.C.8b. Figure A.C.9 finds similar results with NUTS 1-time and NUTS 2-time fixed effects instead of area-time fixed effects. Figure A.C.11b repeats the main analysis on trade shares using the imputation estimator by Borusyak, Jaravel & Spiess (2021). Figure A.C.10 decomposes the intensive margin expansion by sector and Figure A.C.12 by factor intensity. Figure A.C.13 shows the export volume expansion by good. Finaly, Figures A.C.14a and A.C.14b show the export expansion in the long run.

Figure A.C.7: Robustness results for Fact #2: Export volumes and the extensive margin expand



(a) Export volumes



(b) Number of exported goods

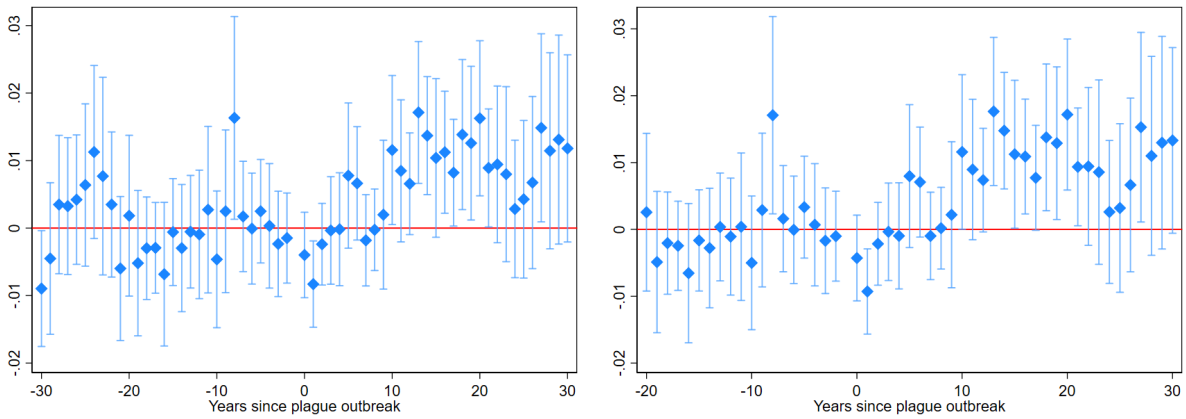
Notes: In Figure A.C.7a, estimation of equation 1.3.2 on $T_{ijt} = x_{ijt}$. Standard errors clustered at the origin level. In Figure A.C.7b, estimation of equation 1.3.2, where T_{ijt} is the number of exported goods.

Table A.C.6: Plague increases exports at the intensive and extensive margins

	Volume		# Goods		# Novel Goods	
	(1)	(2)	(3)	(4)	(5)	(6)
Post Plague	0.503***	38.735***	0.206	0.267***	1.252***	0.090**
	(0.086)	(14.154)	(0.146)	(0.072)	(0.291)	(0.044)
Fixed Effects:						
– Area x Year	✓	✓	✓	✓	✓	✓
– Origin x Destination	✓	✓	✓	✓	✓	✓
– Destination x Year	✓	✓	✓	✓	✓	✓
Estimator	PPML	OLS	PPML	OLS	PPML	OLS
Observations	103,532	361,059	246,958	361,059	46,727	361,059

Notes: Difference-in-differences set up dropping the war years of 1700 to 1721. Standard errors clustered at the origin level in parentheses. * p <0.10, ** p <0.05, *** p <0.01. The dependent variables are annual exports, the number of exported goods, and the number of novel exported goods. The independent variable is a post plague dummy, equal to one after the origin suffered a plague outbreak. Annual growing season temperatures are an additional control in all specifications.

Figure A.C.8: Robustness results for Fact #2: Thirty pre periods and underlying value

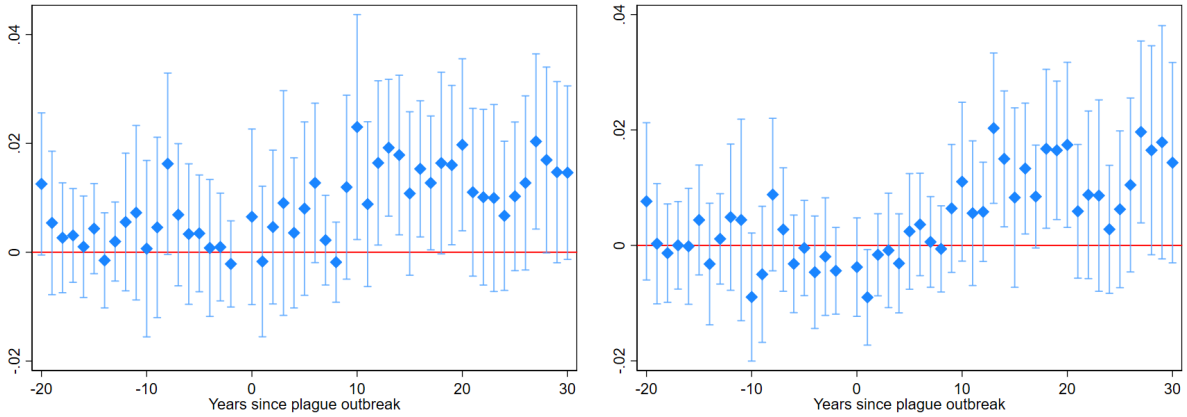


(a) Market shares with 30 pre periods

(b) Market shares, underlying value

Notes: Figure A.C.8a estimates equation 1.3.2 on shares, extending the pre-period to 30 years. Figure A.C.8b shows regression results based on underlying value, as recovered in Appendix A.A.4. Standard errors clustered at the origin level.

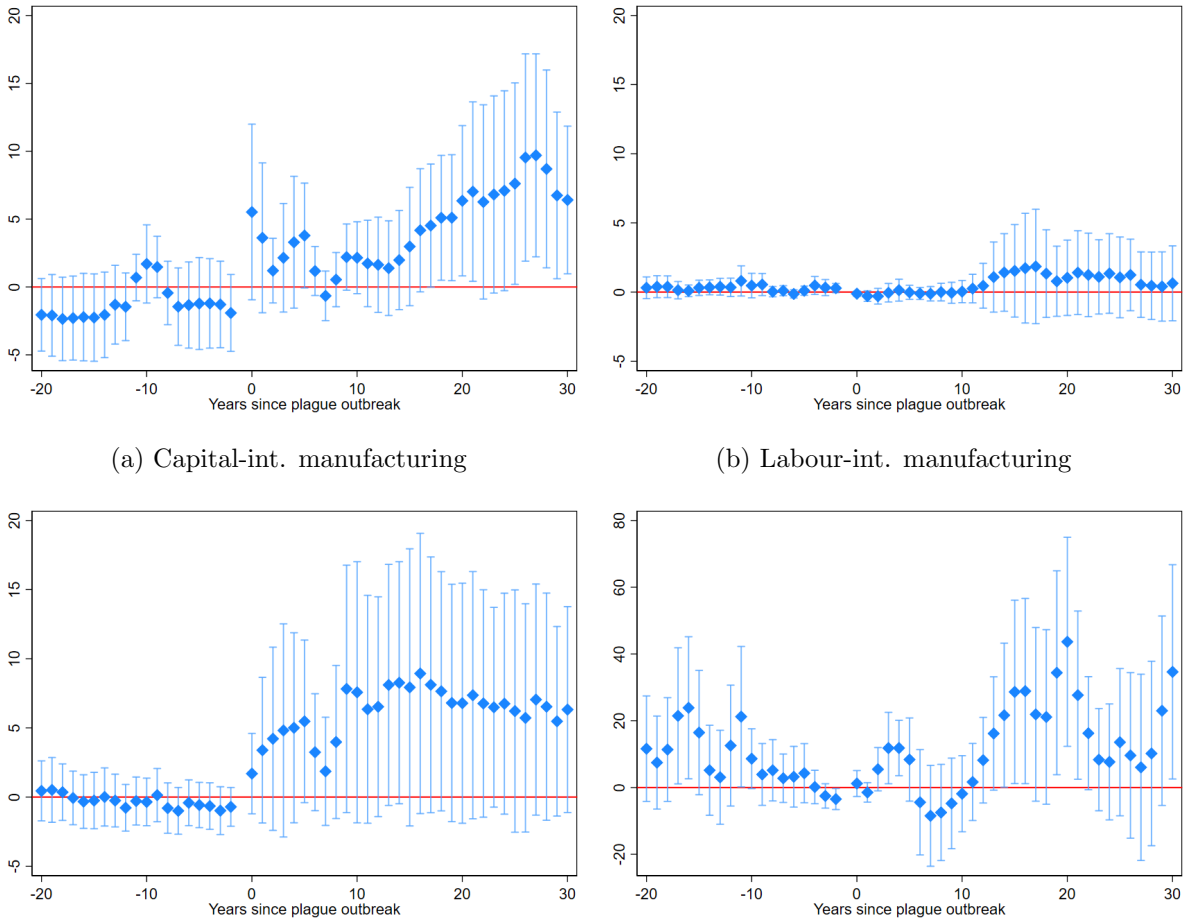
Figure A.C.9: Robustness results for Fact #2: NUTS 1-time and NUTS 2-time fixed effects



(a) Market shares with NUTS 1-time fixed effects (b) Market shares with NUTS 2-time fixed effects

Notes: These Figures estimate equation 1.3.2 on shares. Instead of area-time fixed effects, I use NUTS 1-time and NUTS 2-time fixed effects consistently. Standard errors clustered at the origin level.

Figure A.C.10: Intensive margin expansion, volumes by sector

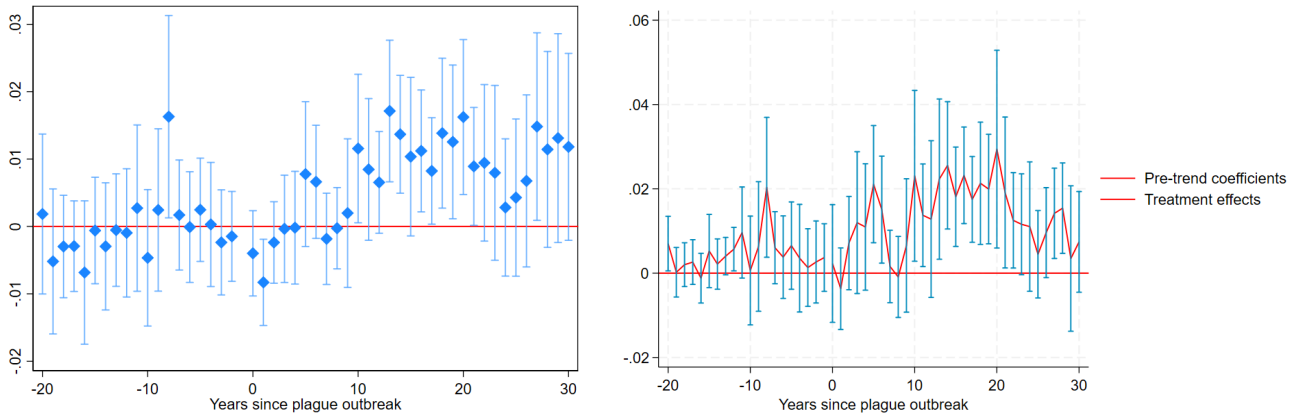


(a) Capital-int. manufacturing (b) Labour-int. manufacturing

(c) Capital-int. agriculture (d) Labour-int. agriculture

Notes: Estimation of equation 1.3.2 on volumes. The y-axis differs in Figure A.C.10d to preserve the legibility of the other Figures. Between 1668-1750, 55% of traded value was in labour-intensive agriculture, 27% in capital-intensive agriculture, 3% in labour-intensive manufacturing and 13 % in capital-intensive manufacturing. Standard errors clustered at the origin level.

Figure A.C.11: Lagged temperature and imputation estimator in Borusyak, Jaravel & Spiess (2021)

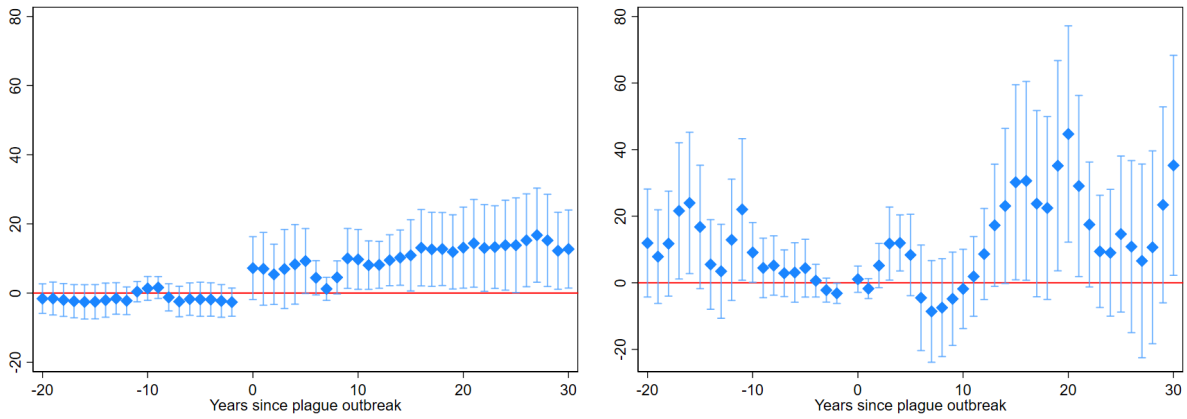


(a) Market shares, adding temperature lags of 10 and 20 years

(b) Market shares, imputation estimator

Notes: Figure A.C.11b shows results from estimating equation 1.3.2 with 20 pre-trends and 30 post periods using the imputation estimator by Borusyak, Jaravel & Spiess (2021). Standard errors clustered at the origin level.

Figure A.C.12: Intensive margin expansion, volumes by factor intensity

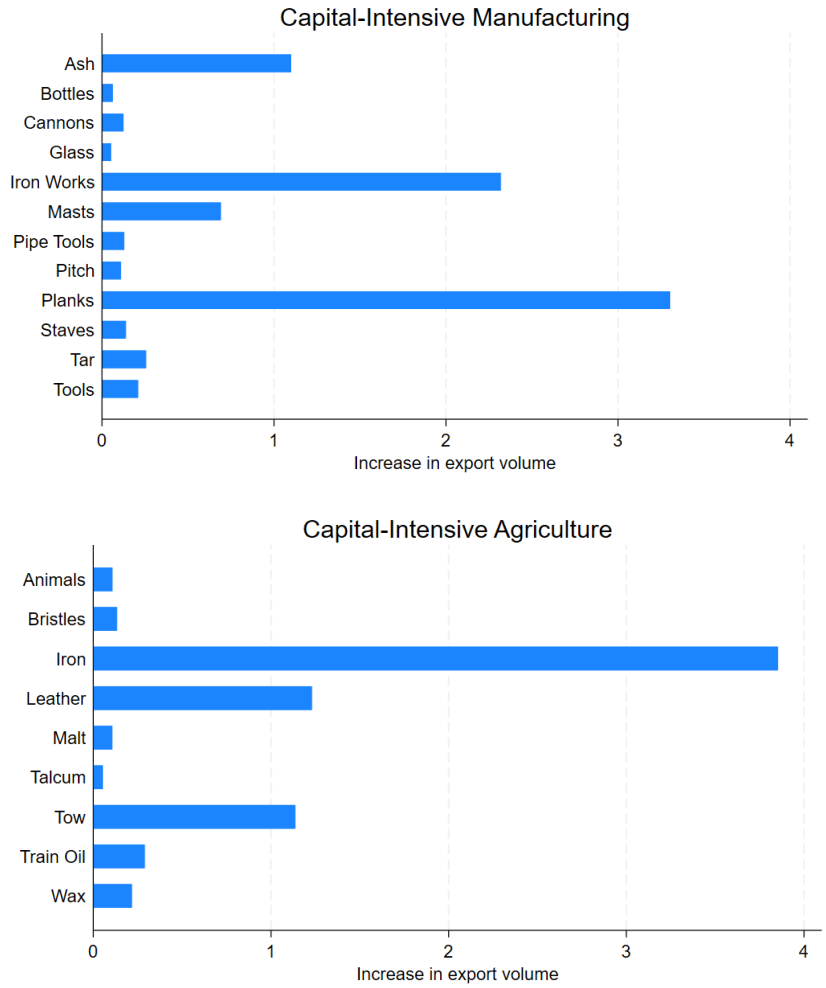


(a) Capital-intensive sectors

(b) Labour-intensive sectors

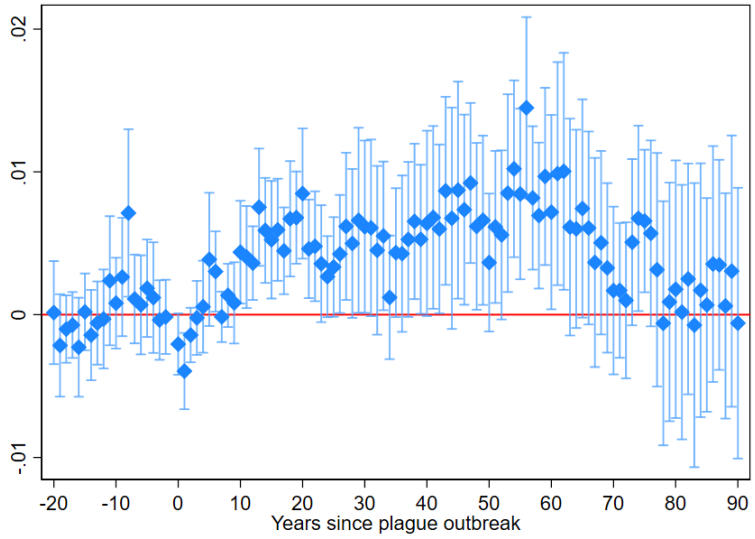
Notes: Estimation of equation 1.3.2 on volumes. Between 1668-1750, 58% of traded value was in labour-intensive sectors and 40% in capital-intensive sectors. Standard errors clustered at the origin level.

Figure A.C.13: Fact #2: Volume expansion by good, capital-intensive sectors

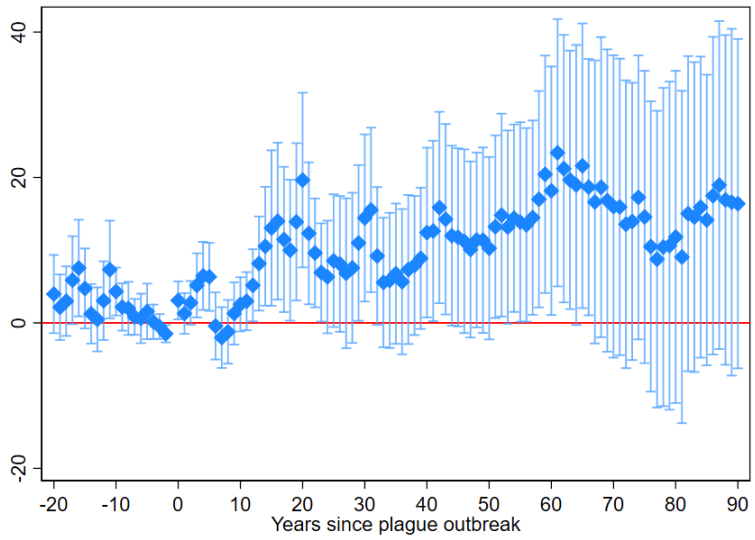


Notes: Results from regressing a plague dummy on i's exports to j in good g in annual data, controlling for area x year, origin x destination, and destination x year fixed effects. Shown are only point estimates that are significant at the 5% level.

Figure A.C.14: Market shares and export volumes expand in the long run



(a) Market shares



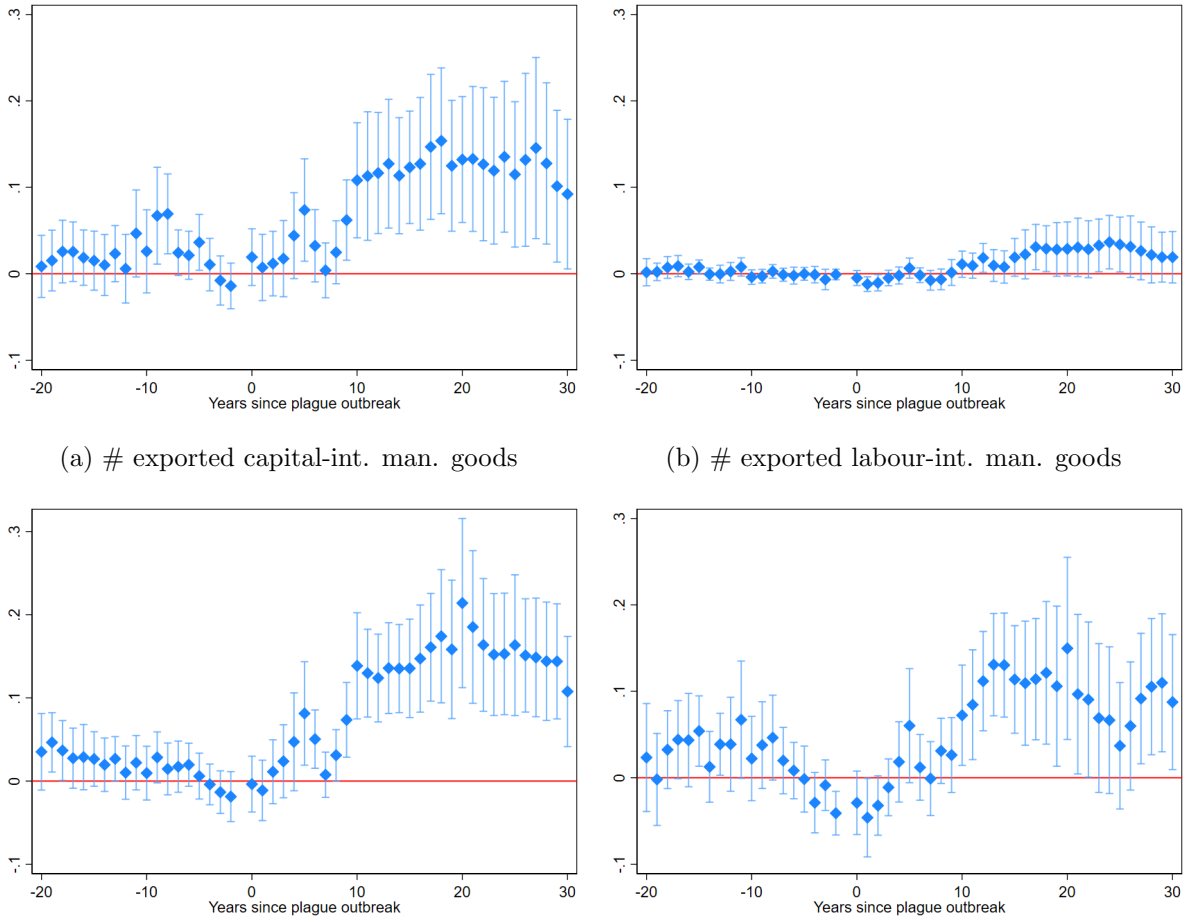
(b) Export volumes

Notes: In Figure A.C.14a, estimation of equation 1.3.2 on i's share of j's imports, $T_{ijt} = \frac{x_{ijt}}{\Sigma_{i \in I} x_{ijt}}$. In Figure A.C.14b, estimation of equation 1.3.2 on $T_{ijt} = x_{ijt}$. Standard errors clustered at the origin level.

A.C.3 Extensive Margin: Plagued regions export a larger variety of goods.

Robustness results for the extensive margin expansion follow. Figure A.C.15 shows that all four sectors participated in the expansion, with labour-intensive manufacturing contributing the least. Figure A.C.16 separates the extensive margin expansion by factor intensity. Figure A.C.17 shows results at the goods level.

Figure A.C.15: Extensive margin expansion, by sector



(a) # exported capital-int. man. goods

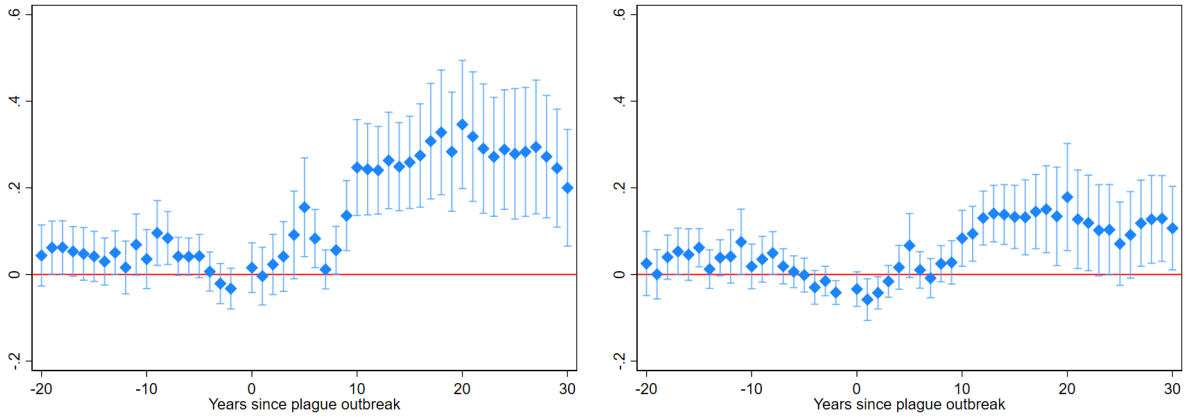
(b) # exported labour-int. man. goods

(c) # exported capital-int. ag. goods

(d) # exported labour-int. ag. goods

Notes: Estimation of equation 1.3.2 on the number of exported goods by sector and factor intensity. Between 1668-1750, 55% of traded value was in labour-intensive agriculture, 27% in capital-intensive agriculture, 3% in labour-intensive manufacturing and 13 % in capital-intensive manufacturing. Standard errors clustered at the origin level.

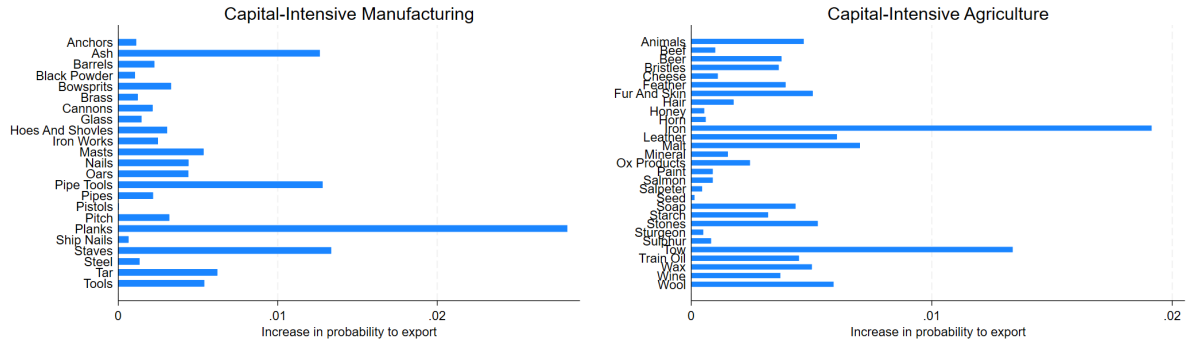
Figure A.C.16: Extensive margin expansion, by factor intensity



(a) # exported capital-intensive goods (b) # exported labour-intensive goods

Notes: Estimation of equation 1.3.2 on the number of exported goods by factor intensity. Between 1668-1750, 58% of traded value was in labour-intensive sectors and 40% in capital-intensive sectors. Standard errors clustered at the origin level.

Figure A.C.17: Fact #3: Extensive margin expansion by good, capital-intensive sectors

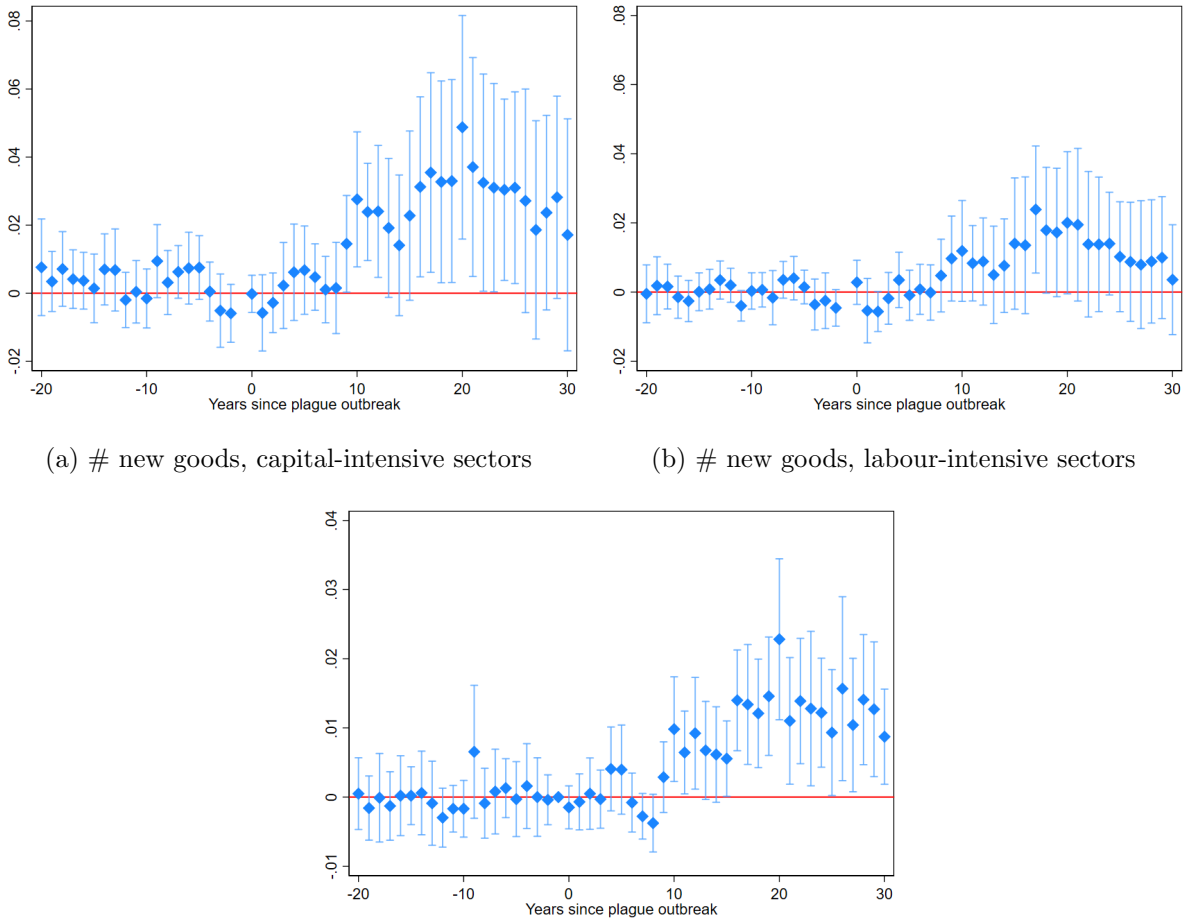


Notes: Results from regressing a plague dummy on a dummy for positive exports of good g from i to j in annual data, controlling for area x year, origin x destination, and destination x year fixed effects. Shown are only point estimates that are significant at the 5% level.

Innovation: Exports of new products increase after the plague.

Figure A.C.18 shows the extensive margin of new goods by factor intensity, while Figure A.C.19 shows these results by good. Table A.C.7 shows results from PPML panel regressions of a post plague dummy on the number of goods never exported before for each of the four sectors. Table A.C.8 shows these results for the volume of goods never exported before, thus illustrating the contribution of new exports to the intensive margin results.

Figure A.C.18: New exports at the extensive margin, by factor intensity



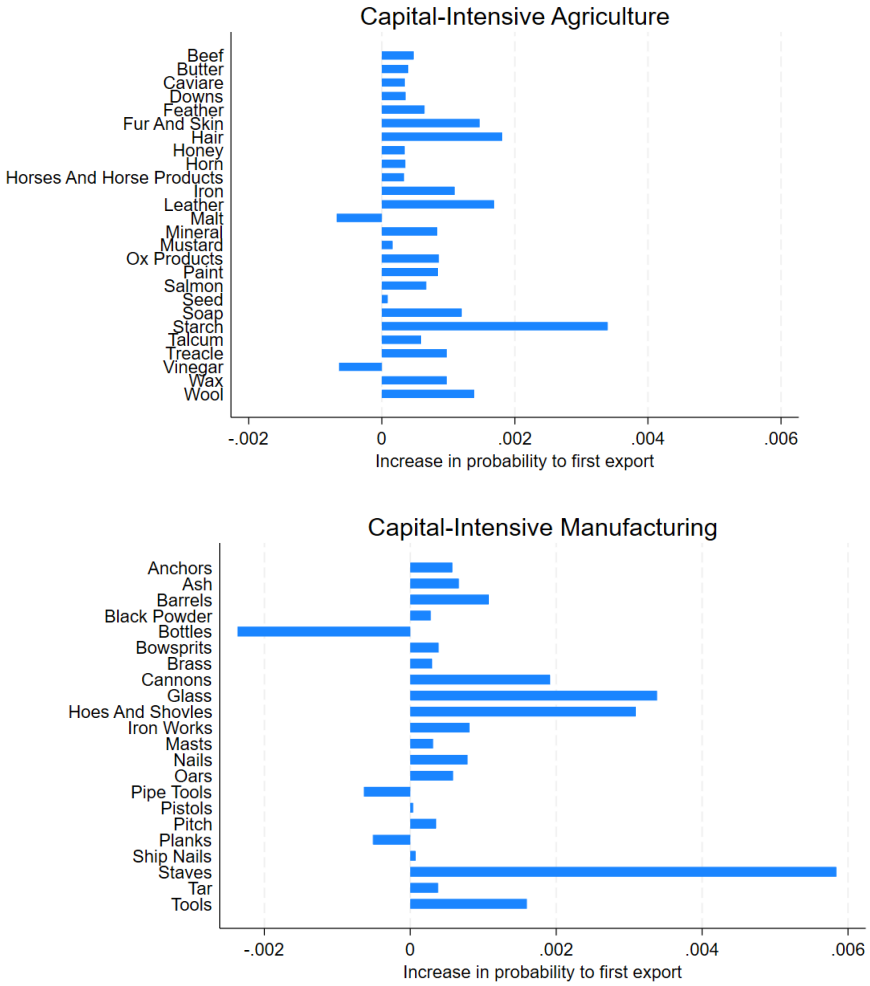
(a) # new goods, capital-intensive sectors

(b) # new goods, labour-intensive sectors

(c) Difference between capital- and labour-intensive sectors

Notes: Estimation of equation 1.3.2 on the number of exported goods, by factor intensity. A new export is defined as a good first exported by a city between 1709, the onset of the plague in my study area, and 1732, 20 years after the last plague year in this region. Standard errors clustered at the origin level.

Figure A.C.19: Fact #4: Increased probability of exporting a good for the first time, capital-intensive sectors



Notes: Results from regressing a plague dummy on a dummy for a new export of good g from i to j in annual data, controlling for area x year, origin x destination, and destination x year fixed effects. A new export is a good not exported before the plague. Shown are only point estimates that are significant at the 5% level.

Table A.C.7: Impact of plague on number of goods never exported before, by factor intensity

	Overall	Labour-intensive ag.	Capital-intensive ag.	Labour-intensive man.	Capital-intensive man.
	(1)	(2)	(3)	(4)	(5)
Post Plague	0.962***	1.011***	0.413	1.349***	0.919***
	(0.242)	(0.325)	(0.280)	(0.266)	(0.321)
Fixed Effects:					
– Area x Year	✓	✓	✓	✓	✓
– Origin x Destination	✓	✓	✓	✓	✓
– Destination x Year	✓	✓	✓	✓	✓
– Controls	✓	✓	✓	✓	✓
Estimator	PPML	PPML	PPML	PPML	PPML
Observations	74,452	17,270	40,554	4,140	14,185

Notes: Standard errors clustered at the origin level in parentheses. * p <0.10, ** p <0.05, *** p <0.01. The dependent variable is the number of goods that a region has never exported before. The independent variable is a plague dummy, equal to one after the origin suffered a plague outbreak. The first 20 years of the data are dropped, as mechanically initially every good is new. Annual growing season temperatures are an additional control.

Table A.C.8: Impact of plague on volume of goods never exported before, by factor intensity

	Overall	Labour-intensive ag.	Capital-intensive ag.	Labour-intensive man.	Capital-intensive man.
	(1)	(2)	(3)	(4)	(5)
Post Plague	1.583***	2.106***	1.260*	1.038**	1.217
	(0.467)	(0.624)	(0.749)	(0.522)	(0.853)
Fixed Effects:					
– Area x Year	✓	✓	✓	✓	✓
– Origin x Destination	✓	✓	✓	✓	✓
– Destination x Year	✓	✓	✓	✓	✓
– Controls	✓	✓	✓	✓	✓
Estimator	PPML	PPML	PPML	PPML	PPML
Observations	50,472	13,365	30,255	3,284	12,149

Notes: Standard errors clustered at the origin level in parentheses. * p <0.10, ** p <0.05, *** p <0.01. The dependent variable is the volume of goods that a region has never exported before. The independent variable is a plague dummy, equal to one after the origin suffered a plague outbreak. The first 20 years of the data are dropped, as mechanically initially every good is new. Annual growing season temperatures are an additional control.

A.C.4 Other Robustness Results

Results from gravity regressions with origin-time, destination-time, and origin-destination fixed effects are displayed in Table A.C.9. Table A.C.10 shows that the plague increased regions' probability to export. For all four facts, I present heterogeneity results in Table A.C.11. In Table A.C.12, I show that days at sea (a measure of travel time and handling times in harbours) are unrelated to both the plague and to capital intensity. Table A.C.13 shows heterogeneity results

by plagued regions' siege status. Figure A.C.20 presents evidence that my finding is not driven by export hubs whose reduced local consumption pushed more goods into exports. I show that plagued regions expanded their imports, speaking against a local demand contraction. Finally, Table A.C.14 shows in data by Baasch (1929) that Hamburg saw an export expanion and a shift into a higher capital intensity of exports after the plague. This speaks against both mechanical explanations and selection issues.

Table A.C.9: Gravity analysis of plague shock on bilateral trade

	Overall	Agriculture		Manufacturing	
		Labour-intensive	Capital-intensive	Labour-intensive	Capital-intensive
	(1)	(2)	(3)	(4)	(5)
Post Plague	0.674*	0.718**	0.470*	1.232***	-0.602
	(0.372)	(0.356)	(0.279)	(0.454)	(0.682)
Fixed Effects:					
– Origin x Year	✓	✓	✓	✓	✓
– Destination x Year	✓	✓	✓	✓	✓
– Origin x Destination	✓	✓	✓	✓	✓
Estimator	PPML	PPML	PPML	PPML	PPML
Observations	505,415	49,887	69,106	26,141	40,514

Notes: Standard errors clustered at the origin level in parentheses. * p <0.10, ** p <0.05, *** p <0.01. The dependent variable is annual bilateral trade. The independent variable is a bilateral plague dummy, equal to one after either the origin, the destination, or both cities suffered a plague outbreak.

Table A.C.10: Impact of plague on probability to export

	Marginal Effect		Average Marginal Effect			
	(1)	(2)	(3)	(4)	(5)	(6)
Post Plague	0.035***	0.007	0.032***	0.154***	0.057***	0.003
	(0.010)	(0.008)	(0.008)	(0.025)	(0.015)	(0.004)
Fixed Effects:						
– Origin		✓		✓		✓
– Year		✓		✓		✓
Estimator	OLS	OLS	Logit	Logit	Tobit	Tobit
Observations	502,980	502,980	502,980	502,980	502,980	502,980

Notes: Standard errors clustered at the origin level in parentheses. * p <0.10, ** p <0.05, *** p <0.01. The dependent variable is a dummy for whether a region exports in a given year. The independent variable is a plague dummy, equal to one after the origin suffered a plague outbreak. For Tobit, I model the exporter dummy as left truncated at 0. For Logit and Tobit, average marginal effects are displayed. In the balanced trade panel, the share of active exporters in a given year averages 10.4%.

Table A.C.11: Heterogeneity of trade findings by pre-plague export levels

	Export Volume		Export Share		# Exported Goods		# New Exported Goods	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Post Plague	-0.98035050**	-0.57260670	0.01024532***	0.00405861	0.19697625***	0.11466535**	0.06830435**	0.02939703
	(0.41571542)	(0.44376456)	(0.00274101)	(0.00311360)	(0.05694356)	(0.05151306)	(0.03415583)	(0.02897934)
Post Plague x Cap-Int.	0.96971886*	1.14320257***						
	(0.52203372)	(0.41610229)						
Post Plague x Pre-Plague Exports		-0.00000115***		0.00000006***		0.00004342*		0.00002052**
		(0.00000037)		(0.00000001)		(0.00002343)		(0.00000889)
Post Plague x Cap-Int. x Pre-Plague Exports		-0.00000154***						
		(0.00000042)						
Fixed Effects:								
– Origin x Sector	✓	✓						
– Sector x Time	✓	✓						
– Area x Year			✓	✓	✓	✓	✓	✓
– Origin x Destination			✓	✓	✓	✓	✓	✓
– Destination x Year			✓	✓	✓	✓	✓	✓
Estimator	PPML	PPML	OLS	OLS	OLS	OLS	OLS	OLS
Observations	102,636	102,636	491,277	491,277	491,277	491,277	491,277	491,277

Notes: Standard errors clustered at the origin level in parentheses. * p <0.10, ** p <0.05, *** p <0.01. The dependent variable in columns 1-2 is the volume of exports by capital intensity. In columns 3-4, it is the market share origin i captures in j. In columns 5-6, it is the number of exported goods. In columns 7-8, it is the number of new exported goods. The independent variables for all four dependent variables are first a post plague dummy and then additionally its interaction with the level of exports before 1709 (number of exported goods) for columns 1-4 (5-8). In columns 1-2, the independent variables also include a post-plague x capint dummy and its interaction with pre-plague exports.

Table A.C.12: Travel times are unaffected by the plague

	Days at Sea			
	(1)	(2)	(3)	(4)
Value of Cap.-Int. Goods	-0.022	-0.010		
	(0.014)	(0.016)		
Total Value	0.032**	0.031		
	(0.015)	(0.025)		
Post Plague x Value of Cap.-Int. Goods		-0.040		
		(0.035)		
Post Plague x Total Value		0.002		
		(0.031)		
Weight of Cap.-Int. Goods			-0.000	0.000
			(0.000)	(0.000)
Total Weight			0.000	-0.000
			(0.000)	(0.000)
Post Plague x Weight of Cap.-Int. Goods				-0.000
				(0.000)
Post Plague x Total Weight				0.000
				(0.000)
Fixed Effects:				
– Captain	✓	✓	✓	✓
– Origin x Year	✓	✓	✓	✓
Measure	Value	Value	Weight	Weight
Observations	66,241	66,241	66,241	66,241

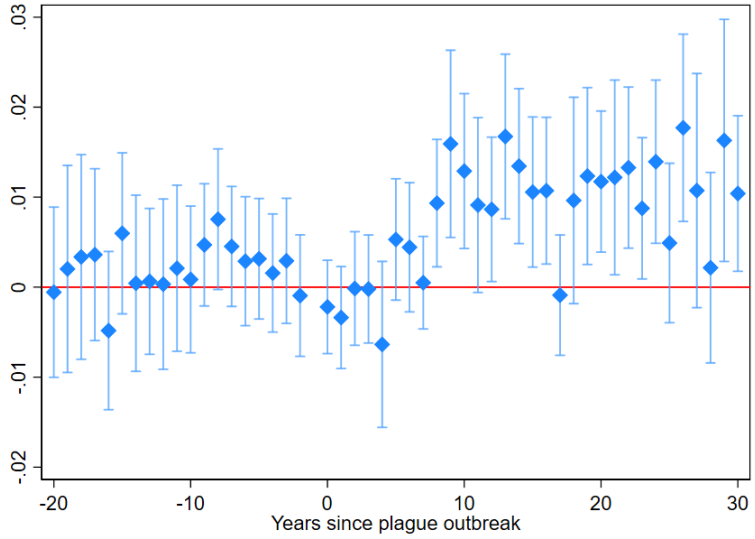
Notes: Standard errors clustered at the origin level in parentheses. * p <0.10, ** p <0.05, *** p <0.01. The sample are trips by the same captain (identified by full name and home town) to and from the Sound, where I restrict to trips where the initial destination is the same as the current origin to avoid including trips with unobserved trips within the Baltic or within the North Sea. To the same end, I exclude observations where more than a year has passed since a captain had last been seen at the Sound. The final sample includes 101,615 passages. Days at sea is the difference in days between two observations at the Sound of the same captain. I control for captain-specific and route-time-specific determinants of travel time with fixed effects.

Table A.C.13: Heterogeneity of trade findings by siege

	Export Volume	Export Share	# Exported Goods	# New Exported Goods
	(1)	(2)	(3)	(4)
Post Plague	-0.948**	0.004	0.128**	0.027
	(0.442)	(0.003)	(0.054)	(0.026)
Post Plague x Pre-Plague Exports	-0.000	0.000***	0.000*	0.000**
	(0.000)	(0.000)	(0.000)	(0.000)
Post Plague x Cap-Int.	1.436***			
	(0.434)			
Post Plague x Cap-Int. x Pre-Plague Exports	-0.000***			
	(0.000)			
Post Plague x Siege	0.471***	0.000	-0.061	0.010
	(0.147)	(0.003)	(0.087)	(0.026)
Post Plague x Cap-Int. x Siege	-0.085			
	(0.228)			
Fixed Effects:				
– Origin x Sector	✓			
– Sector x Time	✓			
– Area x Year		✓	✓	✓
– Origin x Destination		✓	✓	✓
– Destination x Year		✓	✓	✓
Estimator	PPML	OLS	OLS	OLS
Observations	102,636	491,277	491,277	491,277

Notes: Standard errors clustered at the origin level in parentheses. * p <0.10, ** p <0.05, *** p <0.01. The dependent variable in column 1 is the volume of exports by capital intensity. In column 2 it is the market share origin i captures in j. In column 3 it is the number of exported goods. In column 4 it is the number of new exported goods. The independent variables for all four dependent variables are a post plague dummy, its interaction with the level of exports before 1709 (or number of exported goods), and its interaction with a siege dummy. In column 1, the independent variables also includes a post-plague x capint dummy, its interaction with pre-plague exports, and its interaction with a siege dummy. Table A.B.3 shows besieged cities.

Figure A.C.20: Plagued regions also expanded their imports



Notes: Results arguing against mechanical explanations. Standard errors clustered at the level of destinations (as these are import results).

Table A.C.14: Export expansion and shift into capital-intensive goods in post-plague Hamburg

	Absolute		Relative	
	(1)	(2)	(3)	(4)
Post Plague	0.545***	0.439		
	(0.107)	(0.579)		
Post Plague x Capital-Int.			0.569**	0.554
			(0.230)	(0.857)
Fixed Effects:				
– Year			✓	✓
– Sector	✓	✓	✓	✓
Estimator	OLS	PPML	OLS	PPML
Observations	243	304	243	304

Notes: Difference-in-differences set up using Hamburg export data to all destinations between 1702 and 1713, with the plague in 1711, based on Baasch (1929). 38 goods have been classified into sectors precisely as in the rest of the paper. Standard errors in parentheses. * p <0.10, ** p <0.05, *** p <0.01. The dependent variable is annual exports by sector in levels (columns 2 and 4) and logs (columns 1 and 3). The independent variable is a post plague dummy, equal to one after 1711 when Hamburg suffered a plague outbreak, and a dummy interacting the post plague dummy with a dummy for capital-intensive exports. Columns 2 and 4 show unweighted results, whereas in columns 1 and 3 the weight are exports in levels.

A.C.5 Continuous Plague Treatment

As an alternative to a plague dummy, I construct a continuous plague treatment variable, incorporating plague outbreaks across all of Europe and drawing on additional sources for 106 recorded plague outbreaks (Table A.C.15 and Figure A.C.21). A NUTS 3 region counts as plagued if there is at least one plagued town or region in it. The continuous plague treatment is constructed as

$$plague_{it}^{cont} = plague_{it}^{own} + \Sigma_j \frac{Population_j}{Distance_{ij}} plague_{jt},$$

where i are ports, j are regions in Europe at the NUTS 3 level, $_{it}^{own}$ denotes a region's own plague status at time t, and $_{jt}$ is a NUTS 3 region's plague status at time t. In the absence of population data at this resolution, I assume a constant population density and proxy with the size of a NUTS 3 region. Distances are computed as straight lines.[6] As a robustness check, I also construct a treatment that weighs all regions the same by dropping the population (or area) variable:

$$plague_{it}^{cont} = plague_{it}^{own} + \Sigma_j \frac{1}{Distance_{ij}} plague_{jt}.$$

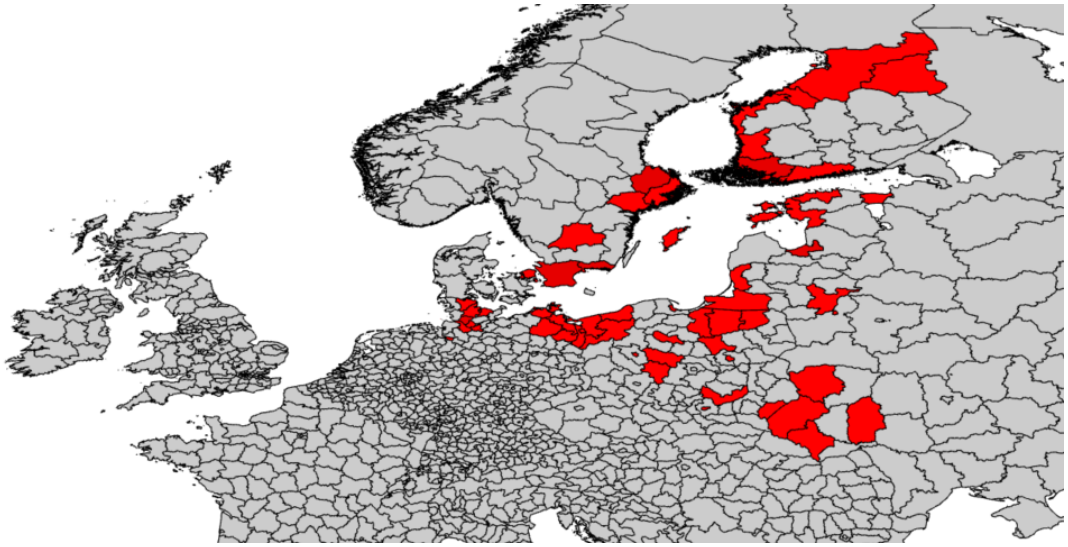
I show in Table A.C.16 that all results hold when regressing on a continuous plague treatment variable. The positive coefficients for the shift into capital-intensive exports after an own plague outbreak is no longer significant, however.

[6] In instances where a region, not a town, is mentioned in the source, I use the coordinates of the historical capital to compute distances.

Table A.C.15: List of plagued cities and regions across all of Europe

Town or Region	Modern Country	Time	Historical Capital	Source
Pińczów	Poland	1702		Frandsen (2009)
Ruthenia	Ukraine	1703–1706	Lwow	Frandsen (2009)
Podolia	Ukraine	1703–1706	Kamianets-Podilskyi	Frandsen (2009)
Volhynia	Ukraine	1703–1706	Lutsk	Frandsen (2009)
Lviv (Lemberg)	Ukraine	1704–1705		Frandsen (2009)
Kołomyja	Ukraine	1705–1706		Jarosław Burchardt et al. (2009)
Stanisławów	Ukraine	1705–1706		Jarosław Burchardt et al. (2009)
Stryj	Ukraine	1705–1706		Jarosław Burchardt et al. (2009)
Sambor	Ukraine	1705–1706		Jarosław Burchardt et al. (2009)
Przemyśl	Poland	1705–1706		Jarosław Burchardt et al. (2009)
Jarosław	Poland	1705–1706		Jarosław Burchardt et al. (2009)
Kraków	Poland	1707		Frandsen (2009)
Lesser Poland	Poland	1707–1710	Krakow	Frandsen (2009)
Mazovia	Poland	1707–1710	Warsaw	Frandsen (2009)
Warsaw	Poland	1707–1710		Frandsen (2009)
Great Poland	Poland	1707–1709	Gniezno	Frandsen (2009)
Ostrów	Poland	1707–1709		Frandsen (2009)
Kalisz	Poland	1707–1709		Frandsen (2009)
Poznań	Poland	1707–1709		Frandsen (2009)
Toruń	Poland	1708		Frandsen (2009)
Piekielko	Poland	1708		Frandsen (2009)
Bialutten	Poland	1708		Frandsen (2009)
Hohenstein	Poland	1708		Sahm (1905)
Masuria	Poland	1708–1710	Elk	Kossert (2005)
Königsberg	Russia	1709–1710		Frandsen (2009)
Danzig (Gdańsk)	Poland	1709		Kroll & Grabinsky (2007)
Elbing (Elblag)	Poland	1709–1710		Frandsen (2009)
Insterburg	Prussia	1708-1711		Kossert (2005)
Memel	Prussia	1708-1711		Kossert (2005)
Ragnit	Prussia	1708-1711		Kossert (2005)
Tilsit	Prussia	1708-1711		Kossert (2005)
Pillupönen (Nevskoye)	Russia	1709		Frandsen (2009)
Stettin (Szczecin)	Poland	1709–1711		Thiede (1849)
Damm (Dabie)	Poland	1709		bei der Wieden (1999)
Pasewalk	Germany	1709–1710		bei der Wieden (1999)
Anklam	Germany	1709–1710		bei der Wieden (1999)
Kammin (Kamień Pomorski)	Poland	1709–1710		bei der Wieden (1999)
Belgard (Białogard)	Poland	1709–1710		bei der Wieden (1999)
Stralsund	Germany	1710–1711		bei der Wieden (1999)
Altentreptow	Germany	1710–1711		bei der Wieden (1999)
Wolgast	Germany	1710–1711		bei der Wieden (1999)
Wollin (Wolin)	Poland	1710–1711		bei der Wieden (1999)
Stargard	Poland	1710–1711		bei der Wieden (1999)
Bahn (Banie)	Poland	1710		bei der Wieden (1999)
Neumark	Poland	1710	Soldin	Schwartz (1901)
Uckermark	Poland	1710	Prenzlau	bei der Wieden (1999)
Prenzlau	Germany	1710		bei der Wieden (1999)
Greifswald	Germany	1711		bei der Wieden (1999)
Lithuania	Lithuania	1709–1713	Vilnius	Frandsen (2009)
Vilnius	Lithuania	1709–1713		Frandsen (2009)
Livonia	Latvia	1709–1711	Riga	Kroll (2006)
Riga	Latvia	1710–1711		Kroll (2006)
Dünamünde (Daugavgrīva)	Latvia	1710		Frandsen (2009)
Estonia	Estonia	1709–1711	Reval	Frandsen (2009)
Reval (Tallinn)	Estonia	1710		Frandsen (2009)
Narva	Estonia	1710–1711		Kroll & Grabinsky (2007)
Pernau (Pärnu)	Estonia	1710		Kroll (2006)
Arensburg (Kuressaare)	Estonia	1710		Frandsen (2009)
Gotland	Sweden	1710–1712	Visby	Bohn (1989)
Visby	Sweden	1710–1712		Bohn (1989)
Stockholm	Sweden	1710–1711		Kroll (2006)
Uppland	Sweden	1710	Stockholm	Persson (2011)
Uppsala	Sweden	1710		Persson (2011)
Södermanland	Sweden	1710	Stockholm	Persson (2011)
Enköping	Sweden	1710		Persson (2011)
Jönköping	Sweden	1710–1711		Kroll & Grabinsky (2007)
Värmdö	Sweden	1711		Frandsen (2009)
Tillinge	Sweden	1711		Frandsen (2009)
Danmark parish	Sweden	1711		Frandsen (2009)
Helsingfors (Helsinki)	Finland	1710		Vourinen (2007)
Borgå (Porvoo)	Finland	1710		Vourinen (2007)
Ekenäs (Tammisaari)	Finland	1710		Vourinen (2007)
Åbo (Turku)	Finland	1710–1711		Vourinen (2007)
Nystad (Uusikaupunki)	Finland	1710–1711		Vourinen (2007)
Raumo (Rauma)	Finland	1710–1711		Vourinen (2007)
Björneborg (Pori)	Finland	1710–1711		Vourinen (2007)
Nädendal (Naantali)	Finland	1710–1711		Vourinen (2007)
Jakobstad (Pietarsaari)	Finland	1710–1711		Vourinen (2007)
Gamlakarleby (Kokkola)	Finland	1710–1711		Vourinen (2007)
Uleåborg (Oulu)	Finland	1710–1711		Vourinen (2007)
Kajana (Kajaani)	Finland	1710–1711		Engström (1994)
Helsingør	Denmark	1710–1711		Frandsen (2009)
Copenhagen	Denmark	1711		Frandsen (2009)
Västana	Sweden	1711		Frandsen (2009)
Karlskrona	Sweden	1710		Persson (2011)
Karlshamn	Sweden	1710		Persson (2011)
Blekinge	Sweden	1710	Karlskrona	Persson (2011)
Domsten	Sweden	1711		Frandsen (2009)
Ystad	Sweden	1712		Frandsen (2009)
Malmö	Sweden	1712		Frandsen (2009)
Hamburg	Germany	1712–1714		Kroll (2006)
Altona	Germany	1713		Frandsen (2009)
Bremen	Germany	1712–1713		Frandsen (2009)
Stade	Germany	1712		Frandsen (2009)
Itzehoe	Germany	1712		Frandsen (2009)
Kropp	Germany	1712		Frandsen (2009)
Glückstadt	Germany	1712		Frandsen (2009)
Rendsburg	Germany	1712		Ulbricht (2004)
Laboe	Germany	1712		Ulbricht (2004)
Holstein	Germany	1712	Glückstadt	Frandsen (2009)
Schleswig	Germany	1712		Ulbricht (2004)
Flensburg	Germany	1712		Ulbricht (2004)
Friedrichsort	Germany	1712		Ulbricht (2004)
Pinneberg	Germany	1712		Winkle (1983)
Rellingen	Germany	1712		Winkle (1983)
Gröpelingen	Germany	1712		Frandsen (2009)

Figure A.C.21: Map of Plagued NUTS 3 Regions



Notes: A map of plagued hinterlands based on Table A.C.15. Red colour indicates 1 or more plagued locations within a NUTS 3 region.

Table A.C.16: Impact of Plague on Trade, Accounting for Outbreaks in the Hinterland

	Plague (continuous treatment)				Plague (only accounting for distance)			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Own Plague	70.19260	0.01037***	0.19819***	0.07497**	68.78410	0.01087***	0.20180***	0.07387**
	(48.95880)	(0.00285)	(0.05970)	(0.03609)	(53.02448)	(0.00295)	(0.06222)	(0.03699)
Plague in Hinterland	-0.02843	-0.00000	-0.00002	-0.00006				
	(0.02638)	(0.00000)	(0.00010)	(0.00004)				
Own Plague x Capital-Int.	1.3e+02				1.6e+02			
	(1.2e+02)				(1.4e+02)			
Plague in Hinterland x Capital-Int.	0.01858*							
	(0.01089)							
Plague in Hinterland, Only Distance					-5.6e+01*	-0.00387**	-0.03760	-0.03447
					(31.34625)	(0.00196)	(0.09991)	(0.02659)
Plague in Hinterland, Only Distance x Capital-Int.					-3.4e+01			
					(49.50306)			
Fixed Effects:								
– Origin x Sector	✓				✓			
– Year	✓				✓			
– Area x Year		✓	✓	✓		✓	✓	✓
– Origin x Destination		✓	✓	✓		✓	✓	✓
– Destination x Year		✓	✓	✓		✓	✓	✓
Estimator	246,095	489,617	489,617	489,617	246,095	489,617	489,617	489,617

Notes: Standard errors clustered at the origin level in parentheses. * p <0.10, ** p <0.05, *** p <0.01. The dependent variable is sectoral exports in columns 1 and 5, trade between i and j in columns 2 and 6, a dummy for positive trade between i and j in columns 3 and 7, and the volume of new export between i and j in columns 4 and 8. The independent variables are an own plague dummy, a continuous plague variable accounting for the size and proximity of other Europe-wide plagued regions, a second such dummy accounting only for proximity of other plagued regions, and their interactions with a capital intensity dummy (only in columns 1 and 5). Annual growing season temperatures are an additional control in all specifications.

Appendix A.D Model Appendix

A.D.1 Recovering Sectoral Productivities

Costinot et al. (2011) show that the following relationship between trade flows, sectoral productivities, and trade costs is true in their model:

$$\frac{X_{ijk}X_{i'jk'}}{X_{ijk'}X_{i'jk}} = \left(\frac{A_{ik}A_{i'k'}}{A_{ik'}A_{i'k}}\right)^{\theta}\left(\frac{d_{ijk}d_{i'jk'}}{d_{ijk'}d_{i'jk}}\right)^{-\theta}. \tag{A.D.1}$$

This equation takes ratios with a reference sector and region. Time variation will be introduced into gravity equation 1.4.8 as follows:

$$X_{ijkt} = \frac{\chi_k A_{ikt}(w_{ikt})^{-\gamma_k\theta}d_{ijkt}^{-\theta}}{CMA_{jkt}}\alpha_k Y_{jt}. \tag{A.D.2}$$

The following result can be established from equation A.D.2:

$$\frac{X_{ijkt}X_{ijk't'}X_{i'jkt'}X_{i'jk't}}{X_{i'jkt}X_{i'jk't'}X_{ijkt'}X_{ijk't}} = \frac{A_{ikt}A_{i'kt'}A_{ik't'}A_{i'k't}}{A_{ikt'}A_{i'kt}A_{ik't}A_{i'k't'}}\left(\frac{w_{ikt}w_{i'kt'}}{w_{i'kt}w_{ikt'}}\right)^{-\gamma_k\theta}\left(\frac{w_{ik't'}w_{i'k't}}{w_{i'k't'}w_{ik't}}\right)^{-\gamma_{k'}\theta}. \tag{A.D.3}$$

I assume that $d_{ijkt} = d_{ijkt'}\forall(ij,kt)$ or that $\frac{d_{ijkt}}{d_{i'jkt}} = \frac{d_{ijkt'}}{d_{i'jkt'}}$, such that trade costs cancel out. Therefore, an orthogonality assumption for trade costs is not required. As I am studying a relatively short period of time before the introduction of the steamship and it seems reasonable to further assume that trade costs were not changed by the plague, I consider this a reasonable assumption.

Three ratios are formed in equation A.D.3. First, a ratio with respect to a reference sector k' to separate productivities from wages. Second, a ratio with respect to a reference origin i', which leads to the demand side cancelling out. Third, a ratio with respect to a reference period t', which means growth relative to a reference period is identified. This three way differencing is precisely what a three-way fixed model delivers (Olden & Møen (2022)). Accordingly, I regress trade flows as follows:

$$X_{ijkt} = exp(\alpha + \delta_{ijt} + \delta_{jkt} + \delta_{ikt}) \times \epsilon_{ijkt}, \tag{A.D.4}$$

where δ_{ijt} is an importer-exporter-time fixed effect, δ_{ikt} an exporter-sector-time fixed effect and δ_{jkt} an importer-sector-time fixed effect. ϵ_{ijkt} is the error term. Note, first of all, that this specification nests all interaction terms of fixed effects and the model is therefore saturated (Angrist & Pischke (2009)). Second, note that the three-way fixed effect δ_{ikt} recovers differences-

in-differences-in-differences (Olden & Møen (2022)).This is to say, δ_{ikt} recovers a time difference of what δ_{ik} recovered in Costinot et al. (2011). Computing equation A.D.3 confirms this finding, such that indeed δ_{ikt} captures:

$$\frac{\delta_{ikt}\delta_{i'kt'}\delta_{ik't'}\delta_{i'k't}}{\delta_{ikt'}\delta_{i'kt}\delta_{ik't}\delta_{i'k't'}} = \frac{A_{ikt}A_{i'kt'}A_{ik't'}A_{i'k't}}{A_{ikt'}A_{i'kt}A_{ik't}A_{i'k't'}}\left(\frac{w_{ikt}w_{i'kt'}}{w_{i'kt}w_{ikt'}}\right)^{-\gamma_k\theta}\left(\frac{w_{ik't'}w_{i'k't}}{w_{i'k't'}w_{ik't}}\right)^{-\gamma_{k'}\theta}. \tag{A.D.5}$$

To arrive at a productivity ratio, two assumptions are required:

1. All sectors have the same labour share, $\gamma_k = \gamma \forall k$.
2. The labour mobility friction does not vary within areas, such that $\phi_{ikt} = \phi_{jkt} \forall (k,t)$ and i,j within the same area.

This allows me to separate wages from productivities. Under serfdom, wages may vary between manufacturing and agriculture. I assume:

$$w_{iMt} = \begin{cases} w_{iAt}, & \text{without serfdom,} \\ (1+\phi_{it})w_{iAt}, & \text{with serfdom.} \end{cases} \tag{A.D.6}$$

Unclassified goods will be the reference sector k', which I assume to produce in the hinterland, paying wage w_{iAt}. Accordingly, for k=LA,CA, I can set $w_{ikt} = w_{ik't} \forall (i,t)$, such that the wage ratio drops out. Thus, δ_{ikt} identifies sectoral productivity growth for agricultural sectors. For manufacturing, k=LM,CM, I plug in wage equalisation assumption A.D.6 to obtain:

$$\delta_{ikt} = \frac{A_{ikt}A_{i'kt'}A_{ik't'}A_{i'k't}}{A_{ikt'}A_{i'kt}A_{ik't}A_{i'k't'}}\left(\frac{(1+\phi_{it})(1+\phi_{i't'})}{(1+\phi_{i't})(1+\phi_{it'})}\right)^{-\gamma\theta}. \tag{A.D.7}$$

The ratio of ϕ parameters needs to be equal to 1 to separate productivities from labour mobility frictions. I assume that time variation in labour mobility and serfdom is constant within areas, such that area-sector-time fixed effects in equation 1.4.9 absorb the ratio of ϕ parameters though serfdom may vary over time.[7] A special case is presented by Denmark (including large parts of modern-day Schleswig Holstein), which is the only country in this area that re-introduced serfdom in 1733 (Gary et al. (2022)). However, as all areas in Denmark were affected by this change, this is inconsequential.

[7]There are two alternative conditions that do not require area-time fixed effects. First, if $\phi_{it} = \phi_{i't}$. At all points in time, the labour mobility friction would have to be the same across regions, which I refuse to assume. A second condition is: if $\phi_{it} = \phi_{it'}$, the ϕ ratio also equals 1. This implies the absence of time variation in the labour mobility friction. Despite evidence for increased serfdom after the plague in Northern Estonia (Raster (2023), the limits to labour mobility may have remained unchanged, and the fundamental hurdle of moving to a German-speaking city while not having been allowed to learn a trade remained in place. In my approach, the plague is allowed to have worsened serfdom; as long as these changes are symmetric within areas, the ϕ ratio will be absorbed by fixed effects.

Section 1.2 outlined how the end of the Little Ice Age saw temperatures rise across Northern Europe. Regions will therefore have seen agricultural productivity growth (Waldinger (2022)). To account for exogenous temperature change, I model agricultural productivity growth as a function of temperature change. In plagued regions, I additionally allow for plague-induced productivity changes. I assume that for a non-plagued region i':

$$\frac{A_{i',At}}{A_{i',At'}} = \Big(\frac{temp_{i',t}}{temp_{i',t'}}\Big)^{\epsilon_A}. \tag{A.D.8}$$

I assume a value of $\epsilon_A = 1$ and correct the recovered agricultural productivity ratios for both labour- and capital-intensive agriculture for this term. This value is reasonable given that most regions are in Northern Europe (see Liu et al. (2020)) and represents an average over estimates for different crops.

A possible concern could be plague-induced reductions in exporter-specific trade costs that are picked up in the δ_{ikt}. However, in Table A.C.12 I show that travel times (the sum of days at sea and handling times at harbours) are unrelated to either the plague or the capital intensity of exports, thus dispelling this concern.

Finally, instead of estimating with OLS as in Costinot et al. (2011), I use PPML on a balanced panel, which is required for the equivalence between their equations (17) and (18). There are a number of reasons for preferring PPML. First of all, the model generates a gravity equation for trade, so all arguments for PPML in Santos-Silva & Tenreyro (2006) apply. The issue of zeros is much more prevalent in my data than in the data by Costinot et al. (2011).

A.D.2 Simulations

Figure A.D.1 shows simulation results comparing simulated productivity growth between the hypothetical full sample and the observed sample. While in the observed sample only trade passing the Sound is registered, such that trade on only one side of the Sound is unobserved, in the theoretical, full sample I simulate trade volumes on all routes.

To simplify matters, I simulate two sectors: a labour-intensive and a capital-intensive one. I choose reference levels for wages, market access, χ_k, α_k, and Y_j and compute trade costs as $d_{ij}^{-\theta}$, where d_{ij} is distance over sea.[8] I simulate two time periods, where in the first all regions have the same capital-intensive sector productivity levels. Plagued regions then experience simulated productivity growth in the capital-intensive sector and have higher levels in the second period,

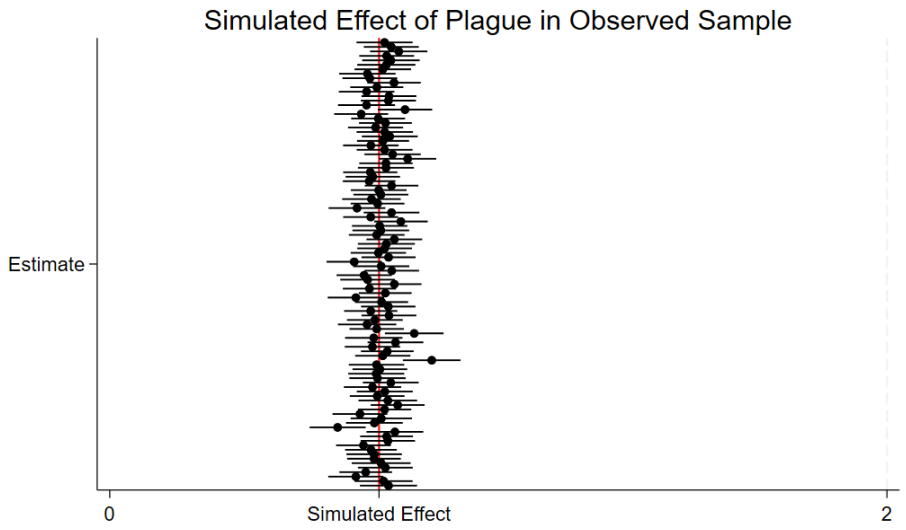
[8] I compute cost distances for each city pair in the dataset using a raster approach and the CostDistance tool in ArcGIS similar to Bakker et al. (2021). My pixel resolution is 30 arc-seconds, corresponding to square cells of about 1 km side length (the longitudinal dimension is even less than that given the latitude of Northern Europe). I compute bilateral distances over sea between all cities. Including land transport is motivated by some large ports slightly inland, such as Thorn or Bordeaux.

whereas non-plagued regions keep their capital-intensive productivity level. The simulated values imply that plagued regions' capital-intensive sector productivity growth is twice as high as that of non-plagued regions. Thus, the simulated effect is $ln2 = 0.693$. Below, I show results for a log-normal specification of noise, where I take the log of simulated trade values, add standard normally distributed noise, and then take the exponential of these noise-amended trade flows. Results with multiplicative and uniformly distributed noise display the same pattern. I then recover sectoral productivities as described in Section A.D.1. I regress a plague dummy on the recovered capital-intensive sector productivity growth and include region and time fixed effects.

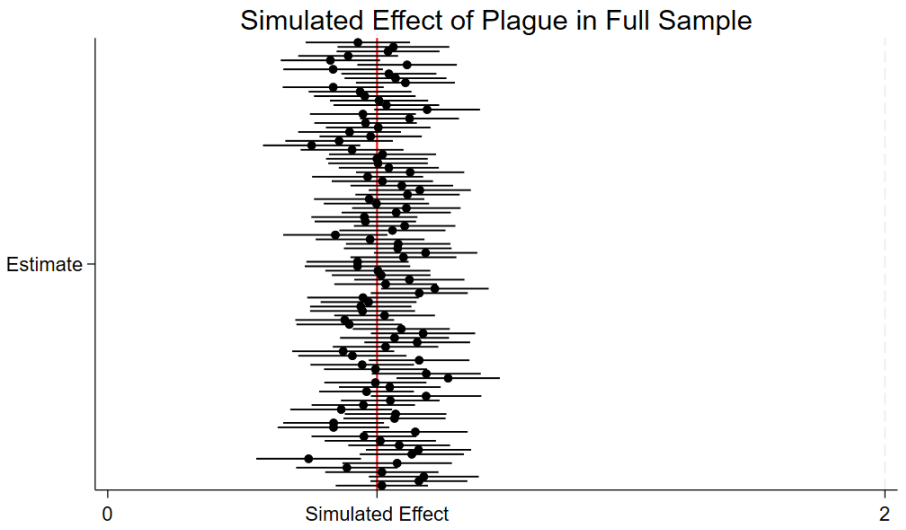
The results are displayed in Figure A.D.1. The simulated effect lies at 0.693, which is within 96 out of 100 confidence intervals for the regressions on the full sample. For the observed sample, 97 out of 100 confidence intervals contain the simulated effect. In all cases, point estimates are centred around the simulated effect.

I also simulate trade following equation 1.4.8 and find similar results when regressing simulated trade volumes in both samples on a plague dummy, including origin-destination and destination-time fixed effects. These results, presented in the bottom panels of Figure A.D.1, show that a significant export increase is picked up in both samples. I run 100 simulations and show that the point estimates are very similar when comparing the full to the observed sample. In both samples, the estimates are centred around the simulated effect, suggesting no bias is introduced by restricting to the observed sample. I conclude that the sample restriction of only observing ships passing through the toll station is not introducing a bias in my productivity growth and trade estimates.

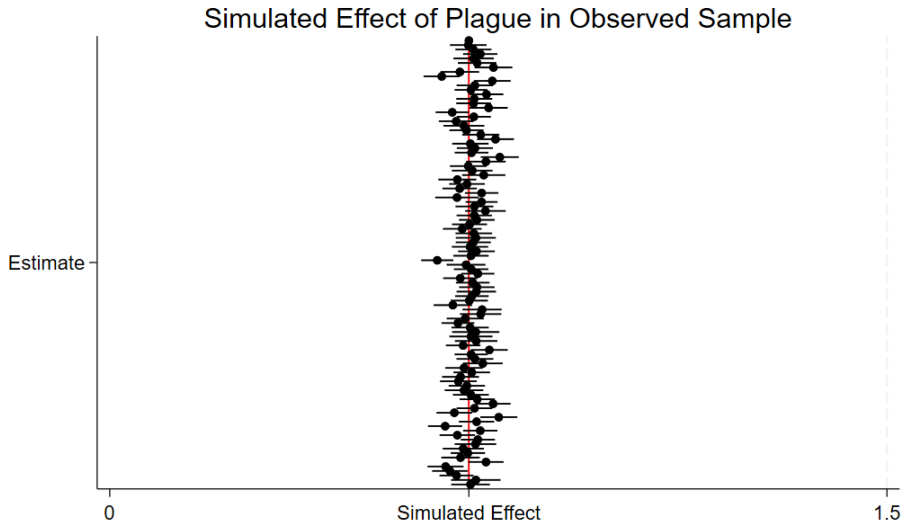
Figure A.D.1: Simulated productivity growth and trade increase in observed vs. full sample



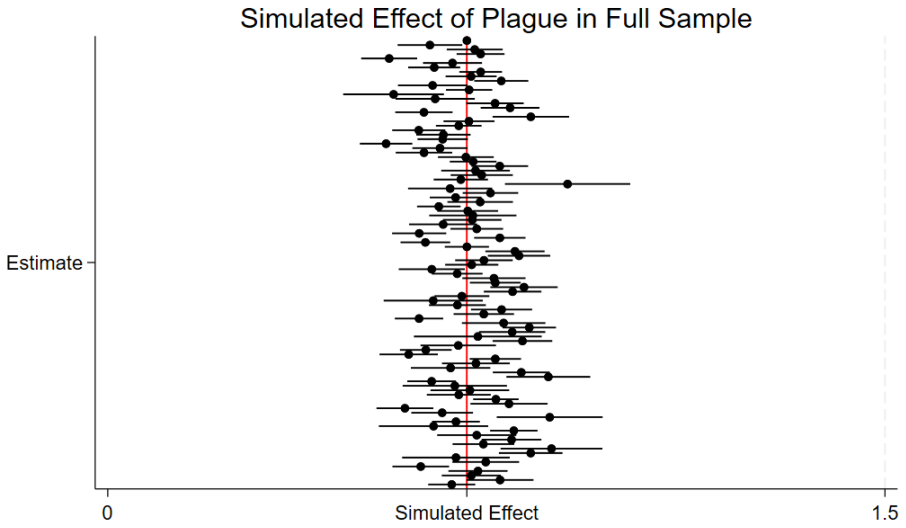
(a) Estimated productivity growth in observed sample



(b) Estimated productivity growth in full sample



(c) Estimated trade growth in observed sample



(d) Estimated trade growth in full sample

Appendix A.E Mechanisms Appendix

Micro-Foundation of Equation 1.5.2

I propose a micro-foundation of equation 1.5.2 that rests on technology adoption choices being affected by changes in relative factor prices. I lean on Bustos (2011), whose framework incorporates technology choice in a Melitz (2003) model. Technology choice is between two technologies: under technology l, fixed costs are lower, but marginal costs are higher. Under technology h, firms need to incur higher fixed costs, yet benefit from lower marginal costs. My model differs from Bustos (2011) in the following ways:

- First, Bustos (2011) is a model of heterogeneous firms. I assume that firms produce varieties within sectors. The firm margin is not defined in my CRS model and thus I never observe firm outcomes. I will instead speak of producers.

- Second, Bustos (2011) features only labour. I add capital as a factor of production. I make the choice between factor intensities discrete at the producer-level: producers choose between a labour- and a capital-intensive technology. This aggregates up to the sectoral capital-to-labour ratio, which remains continuous.

- Third, I turn the model into one of sectoral productivity variation. There is an exogenous component to producers' productivity, which is drawn from distribution $F_{ik}(z)$. Firms can choose to enhance their productivity by investing to operate under the capital-intensive technology, which increases their productivity multiplicatively by υ_k. These productivity-enhancing effects are larger in capital-intensive sectors, $\upsilon_{k'} > \upsilon_k$ if k' is capital- and k labour-intensive.

- Fourth, I model fixed costs to be in units of capital and assume that the labour-intensive technology requires fixed cost if, whereas the capital-intensive technology requires fixed cost $i\rho f$, with $\rho > 1$. This fixed cost has to be paid every period.

I adjust the notation to my setting, where $B_{ik}(\omega)$ replaces ϕ as the exogenous productivity draw. B_{ik} will denote the scale parameter of the Fréchet distribution, which is called A_{ik} in the rest of the paper. In this microfoundation, A_{ik} will instead refer to the realised sectoral productivity incorporating both the exogenous and the endogenous components. This is to be consistent with A_{ik} denoting the measure of sectoral productivity that determines production. υ_k replaces γ as the productivity advantage of the capital-intensive production method, and ρ replaces η as the fixed cost difference.

Firms choose between labour-intensive technology l and capital-intensive technology h. Thus, producers can increase their exogenous productivity draw through endogenous technology choice. Below, I drop the sector-origin subscript for legibility. The production functions are:

$$q_l(\omega) = B(\omega)L_l^{\gamma_l}K_l^{1-\gamma_l}$$
$$q_h(\omega) = \upsilon B(\omega)L_h^{\gamma_h}K_h^{1-\gamma_h},$$

where $\gamma_l > \gamma_h$ and $\upsilon > 1$. The marginal cost functions are:

$$MC_l = \frac{w^{\gamma_l}r^{1-\gamma_l}}{B(\omega)}$$
$$MC_h = \frac{w^{\gamma_h}r^{1-\gamma_h}}{\upsilon B(\omega)}.$$

Producers adopt the capital-intensive technology if $(p - MC_h)q - r\rho f \geq (p - MC_l)q - rf$, where p is the price and r is the return rate of capital. This can be re-arranged to:

$$q \geq \frac{rf(\rho - 1)}{MC_l - MC_h}. \tag{A.E.1}$$

The marginal cost difference is given by:

$$MC_l - MC_h = \frac{\upsilon w^{\gamma_l}r^{1-\gamma_l} - w^{\gamma_h}r^{1-\gamma_h}}{\upsilon B(\omega)}. \tag{A.E.2}$$

Equation A.E.2 is increasing in the factor price ratio $\frac{w}{r}$. The sign of the fraction equals the sign of the numerator, and as $\gamma_l > \gamma_h$ the numerator's first term grows faster in $\frac{w}{r}$ than the second term. Therefore, $\frac{\partial(MC_l - MC_h)}{\partial(\frac{w}{r})} > 0$. In sum, as wages rise relative to the price of capital, the relative marginal cost of the labour-intensive technology rises.

As a result, the cut-off productivity $B_{ik}^*(\omega)$ drops. To see this, note that quantity is increasing in productivity as per the production functions. Given equation A.E.1, the cutoff quantity q^* drops as relative wages rise. As quantity increases with productivity, the productivity threshold $B_{ik}(\omega)^*$ also falls. More producers now adopt the capital-intensive technology. These are the most productive producers among those that have previously used labour-intensive technology. This aggregates up to sectoral productivity, now including both the exogenous and the endogenous productivity components, where I now re-introduce the origin-sector subscript:

$$A_{ik} = \int_{B_{ik}^{min}}^{B_{ik}^*(\omega)} B_{ik}(\omega)\, dF_{ik} + \int_{B_{ik}^*(\omega)}^{B_{ik}^{max}} \upsilon_k B_{ik}(\omega)\, dF_{ik}. \tag{A.E.3}$$

A_{ik} is increasing in $\frac{w}{r}$ as $B_{ik}^*(\omega)$ is decreasing in $\frac{w}{r}$: as relative wages rise, more producers adopt the capital-intensive technology, and are thus shifted into the second term of the sum,

which increases the left-hand side. Similarly, the capital-to-labour ratio increases in $\frac{w}{i}$, which can be seen by aggregating producers' first order conditions:

$$\frac{K_{ik}}{L_{ik}} = \frac{\int_\omega K_{ik}(\omega)}{\int_\omega L_{ik}(\omega)}. \tag{A.E.4}$$

As both A_{ik} and $\frac{K_{ik}}{L_{ik}}$ are strictly increasing in $\frac{w}{r}$ and invertible, there exists a function:

$$A_{ik} = C\left(\frac{K_{ik}}{L_{ik}}\right)^{\beta_k}, \tag{A.E.5}$$

where constant C will be absorbed by fixed effects. Thus, equation A.E.5 corresponds to equation 1.5.2. β_k captures several channels: the direct productivity-enhancing effect of technology adoption within producers, which varies by sector as υ_k is larger in capital-intensive sectors; the fact that the most productive producers using the labour-intensive technology switch to the capital-intensive one; and the reallocation effect, as market shares are gained by more productive producers.

While I do not observe firm or producer outcomes, there is empirical evidence in support of this mechanism. Given fixed exporting costs in this framework, higher productivity from technology adoption leads more producers to profitably begin exporting. The extensive margin expansion documented in Figure A.C.7b may reflect this producer entry into exporting. This involves both new producers exporting goods previously exported by others and producers exporting novel goods. Figure A.E.6 shows that exports of such novel export goods, never exported before by a region, expand after the plague.

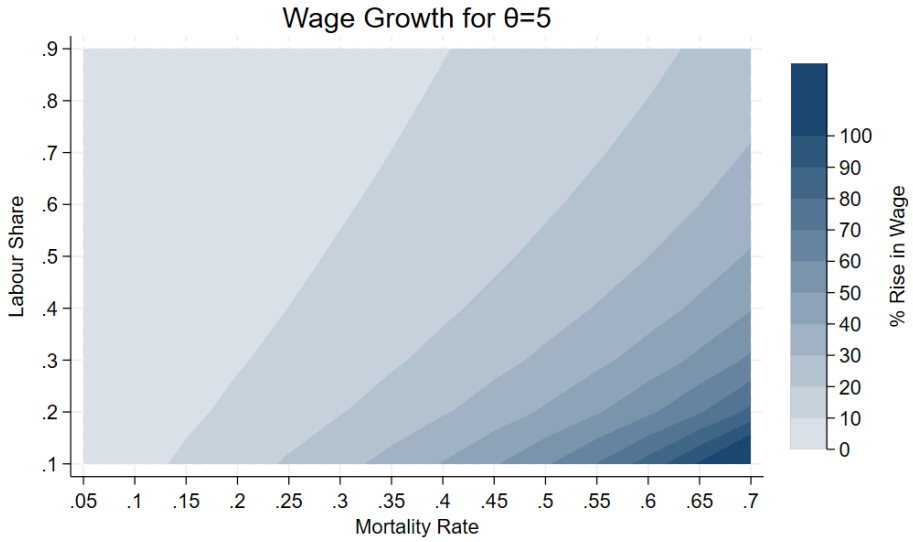
The extensive margin expansion and addition of novel export goods are further stronger in capital-intensive sectors (Appendix Figures A.C.16 and A.C.18c). When interpreting this as more producers beginning to export, it implies that technology adoption leads in particular more producers in capital-intensive sectors to begin exporting. Given my assumptions on υ_k being larger in those sectors, the productivity-enhancing effects of technology adoption and thus the shift into exporting are indeed larger in capital-intensive sectors. This is again consistent with the above mechanism.

Step #1: The Plague Induces Labour Scarcity

Figure A.E.1 shows model-implied wage growth for the entire observed range of mortality rates and labour shares between 0.1 and 0.9. Figure A.E.2 shows an event study for the number of captains living in a region before and after the plague. Table A.E.1 shows that grain price-inferred

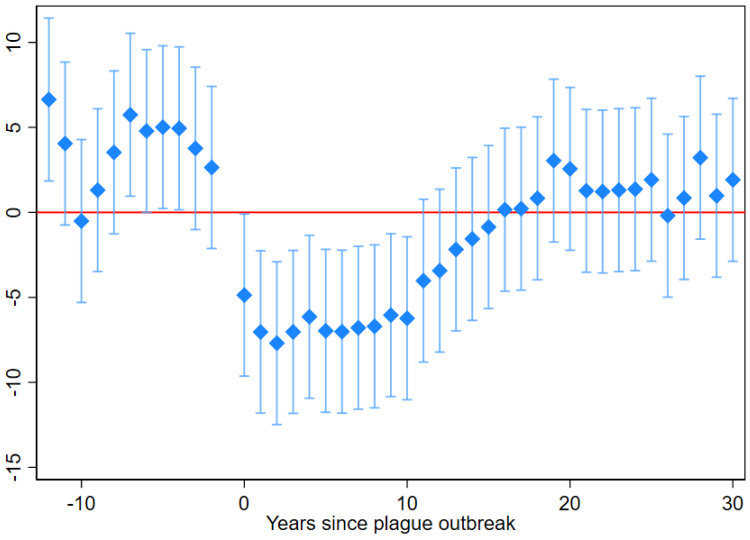
interest rates in Danzig did not rise after the plague.

Figure A.E.1: Model-implied wage growth by mortality rate and labour share



Notes: This figure shows model-implied wage growth for different mortality rates and labour shares.

Figure A.E.2: Dynamics of population recovery after the plague: number of captains as proxy for population



Notes: Event study with area-year and region fixed effects on the number of unique captains' last names living in a region. Captains with the same first and last name from the same region in the same year are assumed to be duplicates and thus dropped.

Table A.E.1: Grain price-inferred interest rates in Danzig do not increase

	Wheat		Rye		Barley	
	(1)	(2)	(3)	(4)	(5)	(6)
Post Plague	-0.042	0.006	-0.065**	0.025	0.055	0.013
	(0.036)	(0.024)	(0.030)	(0.025)	(0.053)	(0.018)
Sample	Full	r<0.1	Full	r<0.1	Full	r<0.1
Observations	197	160	259	199	172	116

Notes: Standard errors in parentheses. * p <0.10, ** p <0.05, *** p <0.01. The dependent variables are grain price-inferred monthly interest rates in Danzig between 1704 and 1793, using the methodology introduced by Keller et al. (2021). I assume that storage costs are constant. Grain price data by Allen & Unger (2018). The independent variable is a post-plague dummy that takes the value 1 from 1709 onwards, when Danzig had been plagued. In some specifications, inferred interest rates above 10% have been dropped, as some underlying grain prices were distorted by outliers during war years.

A.E.1 Step #2: Production Becomes More Capital-Intensive

Table A.E.2 lists examples of capital deepening by sector in this period and area. Table A.E.3 shows that following a plague outbreak, ports significantly increased the number of ships they owned, which I interpret as a proxy for the capital stock. Figure A.E.3 instead presents an event study on this proxy. Figure A.E.4 shows that imports of capital goods increase after the plague. The only transiently increasing share of imports fits with the finding of stable relative imports in Section A.E.5 below. Figure A.E.5 shows Scania's overall shift out of arable and into pastoral farming after the plague. Table A.E.4 shows for individual farms that this shift was stronger the closer a farm was to plagued cities. Table A.E.5 shows capital deepening in agriculture after the plague as more horses per capita are employed in plagued Swedish parishes. Table A.E.6 shows that labour scarce farms with rich soils engage in stronger capital deepening. Figure A.E.6 shows an event study for novel exports after the plague.

Table A.E.2: Examples of capital deepening by sector

Panel A: Labour-Intensive Agriculture

Example	**Description**
Iron ploughs	Replacing of wooden ploughs with iron-tipped or full-iron ploughs
Mechanical sowing	Replacing broadcast method, reducing seed waste
Horse-drawn harrows	Shift from manual to horse-drawn harrowing
Drainage infrastructure	Ditching, dyking, and polder construction
Grain storage	Fixed structures reducing post-harvest losses
Scythe	Replacing sickle, raising area harvested per worker
Wind-/watermills	Replacing hand-grinding

Panel B: Capital-Intensive Agriculture

Example	**Description**
Enclosure and fencing	Replacing herding labour with fixed capital
Livestock housing	Purpose-built winter housing extending productive season
Selective breeding	Improved cattle and sheep breeds for dairy/wool yield
Dairy processing	Butter churns and cheese presses raising output per worker
Hay/fodder storage	Hay barns and winter fodder infrastructure raising carrying capacity
Fish processing	Smokehouses, salt stores, drying racks for herring, cod, stockfish

Panel C: Labour-Intensive Manufacturing

Example	**Description**
Dutch ribbon loom	Multi-ribbon loom weaving 12-24 ribbons simultaneously
Stocking frame	Replacing hand-knitting with frame
Fulling mills	Water-powered fulling replacing manual trampling of woollen cloth
Finishing equipment	Mangles, tentering frames, presses for cloth finishing
Bleachfields	Dedicated infrastructure for linen/cloth bleaching
Warping mills	Replacing labour with mechanical preparation of warp threads
Dyeing vats	Dyeing infrastructure with controlled heating

Panel D: Capital-Intensive Manufacturing

Example	**Description**
Coke-fired blast furnaces	Substitution of coke for charcoal in iron smelting (1709)
Larger blast furnaces	Scaling of charcoal furnaces with water-powered bellows
Slitting & rolling mills	Replacing hand-hammering with mechanical processing of bar iron
Water-powered sawmills	Replacing pit-sawing with mechanised timber processing
Dry docks	Fixed infrastructure for ship construction and repair
Shipyard cranes	Treadwheel and horse-powered cranes for heavy lifting
Ropewalks	Purpose-built long structures for rope production
Brick kilns & glassworks	Improved kiln and furnace design

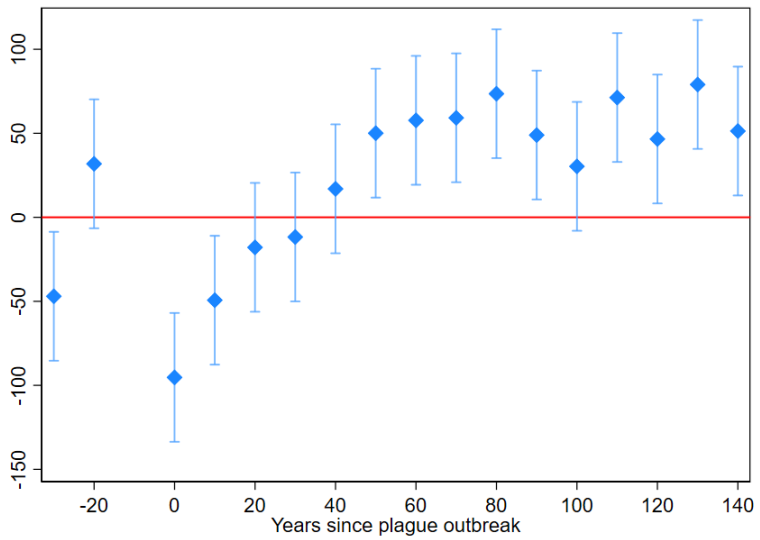
Notes: These are illustrative examples of capital deepening by sector that are plausible for the Northern European and Baltic context around 1700.

Table A.E.3: Impact of plague on proxied capital stock

	# Ships							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Plague Dummy	166.218***	104.535***	164.057***	94.835***				
	(45.161)	(23.508)	(45.329)	(23.547)				
Mortality Rate					385.187***	248.952***	380.036***	226.573***
					(119.599)	(64.118)	(119.693)	(63.366)
Fixed Effects:								
– Region		✓		✓		✓		✓
– Decade			✓	✓			✓	✓
Observations	41,400	41,400	41,400	41,400	41,253	41,253	41,253	41,253

Notes: Standard errors clustered at the origin level in parentheses. * p <0.10, ** p <0.05, *** p <0.01. The dependent variable is the number of ships registered in a city. It is assumed that ships do not sit idle and therefore that every owned ship appears once a decade. Further I assume that ships from the same city with the same captain in the same year are duplicates and drop these. The independent variable is first a plague dummy that equals 1 for plagued cities after the plague hit. The second independent variable is the mortality rate, which for half of regions is imputed as the predicted value from regression results presented in Appendix Table A.B.4.

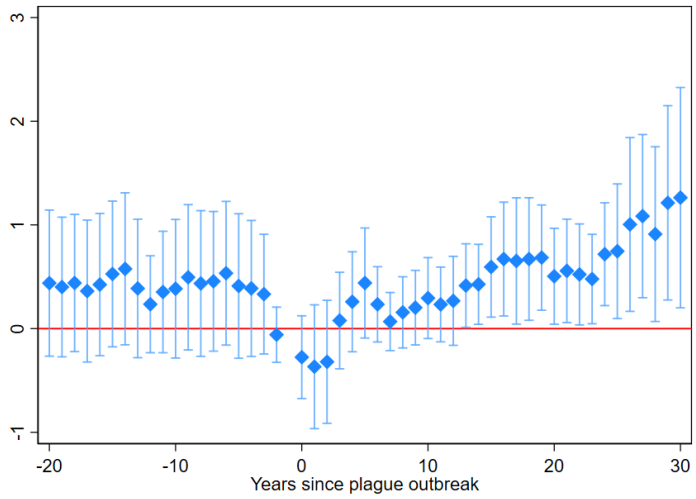
Figure A.E.3: Dynamics of capital accumulation after the plague: number of ships as proxy for the capital stock



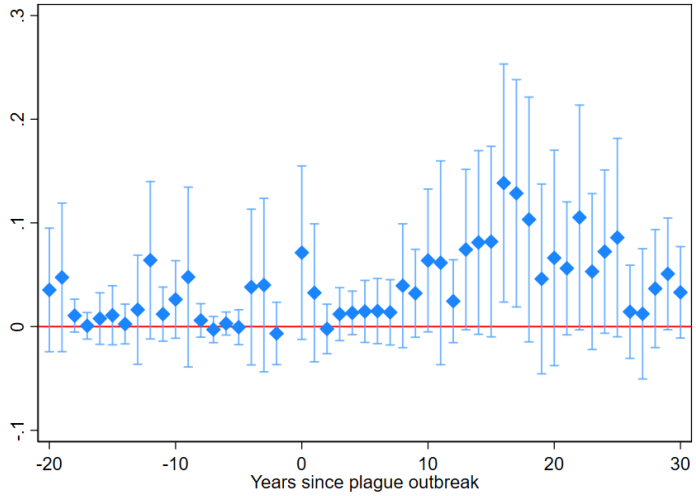
Notes: Event study with area-time, region, and decade fixed effects on the number of ships registered in a harbour in levels. Decade -1 is omitted as the reference decade.

Figure A.E.4: Imports of capital goods increase after the plague, suggesting capital accumulation

(a) Imports of capital goods in levels

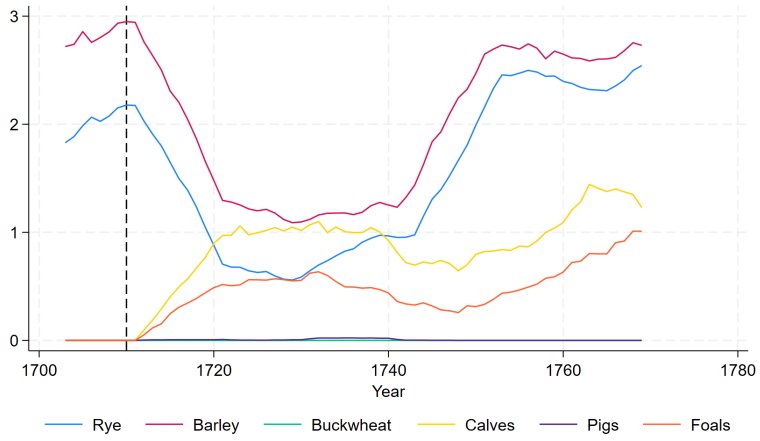


(b) Imports of capital goods as a share of imports



Notes: Capital goods are defined as metals (iron, iron works, steel, lead, copper, tin, sheet metal, brass, brass works), tools (tools, nails), and building materials (planks, tiles, stones, cement, chalk). In panel (a), I regress on imports in levels and include area-time, importer-exporter, and exporter-year fixed effects and control for growing season temperatures. In panel (b), I regress on ports' share of capital goods in imports, including importer and area-time fixed effects and controlling for growing season temperatures. In both panels, standard errors are clustered at the importer level.

Figure A.E.5: Agricultural production in Scania becomes more capital-intensive



Notes: Figure A.E.5 covers 119 Scanian farms and the composition of their production.

Table A.E.4: Farm production and distance to plagued cities

	Calves	Pigs	Foals	Rye	Barley	Buckwheat
	(1)	(2)	(3)	(4)	(5)	(6)
Plague Distance x Post Plague	-0.00656**	-0.000395**	-0.00521***	0.0182***	0.0355***	0.000228**
	(0.00281)	(0.000156)	(0.00171)	(0.00138)	(0.00310)	(0.0000965)
Fixed Effects:						
– Year	✓	✓	✓	✓	✓	✓
– Farm	✓	✓	✓	✓	✓	✓
Observations	5,683	5,683	5,683	5,304	5,304	5,683

Notes: Standard errors clustered at the farm level in parentheses. * p <0.10, ** p <0.05, *** p <0.01. The dependent variables are farm's production of calves, pigs, foals rye, barley, and buckwheat. The independent variable is the sum of distances multiplied by a plague dummy over the four closest cities in Scania province. Both Ystad and Malmö suffered a plague outbreak in 1712 with a mortality rate of 38% and 35%, respectively, whereas Landskrona and Helsingborg did not suffer plague outbreaks.

Table A.E.5: Parish-level shift into higher capital intensity within agriculture

	Horses	Log Horses
	(1)	(2)
Plague & Year=1750	34.22**	0.232***
	(13.71)	(0.0527)
Plague & Year=1810	51.02***	0.334***
	(9.272)	(0.0524)
Population	0.0858***	
	(0.0135)	
Log Population		0.157***
		(0.0532)
Fixed Effects:		
– Parish	✓	✓
– Year	✓	✓
Observations	7,031	5,695

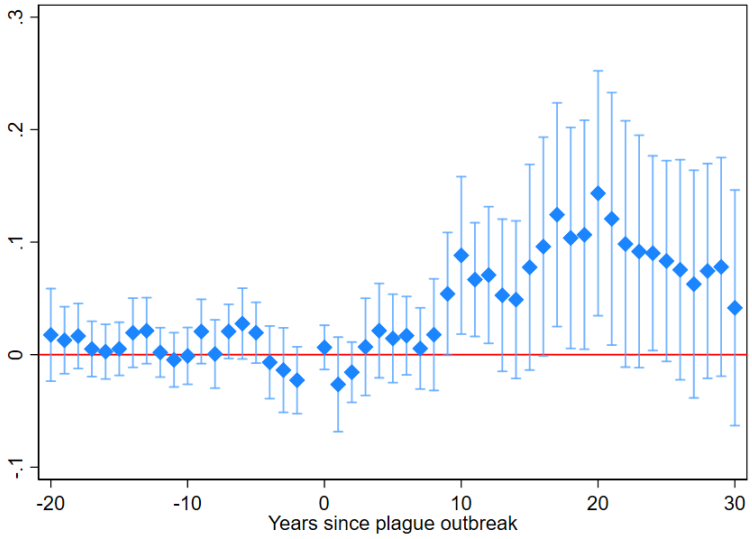
Notes: Standard errors clustered at the parish level in parentheses. * p <0.10, ** p <0.05, *** p <0.01. The dependent variables are parish-level logarithmic stocks of horses in 1690, 1750, and 1810. The independent variables are a post-plague dummy and parish-level population.

Table A.E.6: Farm production, distance to plagued cities, and soil quality

	Calves	Pigs	Foals	Rye	Barley	Buckwheat
	(1)	(2)	(3)	(4)	(5)	(6)
Plague Distance x Post Plague	-0.00598	0.0000460	0.00425	-0.0117	0.0405***	-0.000353
	(0.00924)	(0.000181)	(0.00575)	(0.00894)	(0.0105)	(0.000482)
Soil Quality x Post Plague	0.183	0.0113	0.401**	-1.156***	0.451	-0.00754
	(0.260)	(0.00798)	(0.157)	(0.258)	(0.270)	(0.0130)
Plague Distance x Soil Quality x Post Plague	-0.000356	-0.000124	-0.00284*	0.00913***	-0.00160	0.000155
	(0.00262)	(0.0000845)	(0.00149)	(0.00269)	(0.00293)	(0.000139)
Fixed Effects:						
– Year	✓	✓	✓	✓	✓	✓
– Farm	✓	✓	✓	✓	✓	✓
Observations	5,683	5,683	5,683	5,683	5,683	5,683

Notes: Standard errors clustered at the farm level in parentheses. * p <0.10, ** p <0.05, *** p <0.01. The dependent variables are farm's production of calves, pigs, foals, rye, barley, and buckwheat. The specification interacts a post-plague indicator with (i) plague distance - the sum of distances to the four closest cities in Scania weighted by a plague dummy - (ii) average soil quality - a weighted index over ten soil-quality categories recorded in the land surveys, ranging from 1 (worst) to 10 (best) - and (iii) the product of plague distance and soil quality. Both Ystad and Malmö suffered a plague outbreak in 1712 with a mortality rate of 38% and 35%, respectively, whereas Landskrona and Helsingborg did not suffer plague outbreaks.

Figure A.E.6: New exports at the extensive margin

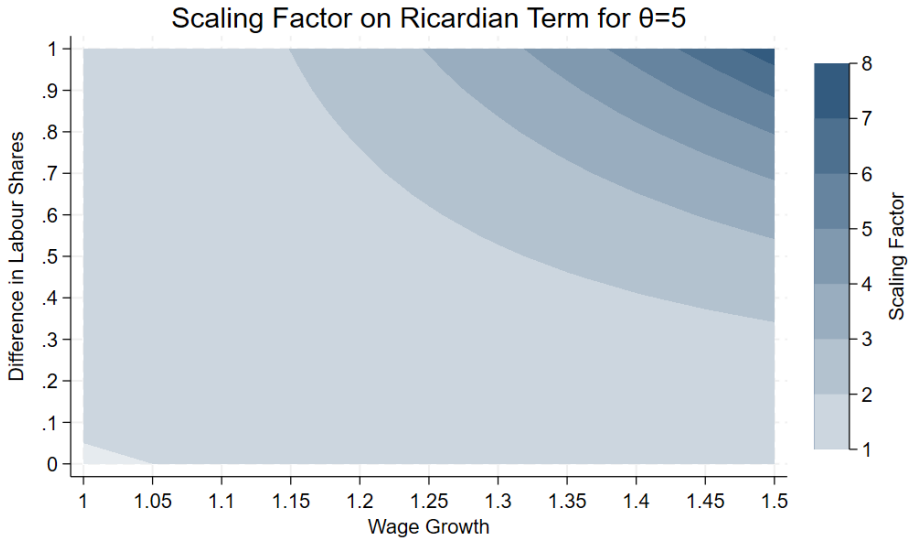


Notes: Estimation of equation 1.3.2 on the number of exported goods. A new export is defined as a good first exported by a city between 1689, 20 years before the onset of the plague in my study area, and 1732, 20 years after the last plague year in this region. Standard errors clustered at the origin level.

A.E.2 Step #3: Productivity Grows More in High β_k Sectors

Figure A.E.7 shows the factor by which the Ricardian term is scaled up in case my assumption on equal labour shares is violated. If $\gamma_k \neq \gamma_{k'}$, δ_{ikt} recovers the product of relative productivity growth and relative wage growth. Specifically, relative productivity growth is multiplied by $\left(\frac{w_{ikt}w_{i'kt'}}{w_{i'kt}w_{ikt'}}\right)^{\theta(\gamma_{k'}-\gamma_k)}$. The second term drops out if $\gamma_k = \gamma_{k'}$; for $\gamma_k < \gamma_{k'}$, the wage ratio leads to an inflation in δ_{ikt} that is assigned to productivity. I put bounds on this potential bias. For model-implied wage growth of 6.6%, $\theta = 5$ and the highest possible $\gamma_k - \gamma_{k'}$ of 1, relative productivity growth is scaled up by 37.6%. Thus, when comparing a sector using only capital to a sector using only labour under the assumption of equal labour shares, the difference in point estimates by factor intensity is 0.57 in manufacturing (from 0.79 without the bias bounds correction) and 0.6 in agriculture (from 0.82). For a more reasonable difference in labour shares of 0.5, the scaling up amounts to 17.3%. Figure A.E.7 shows the full range of this scaling factor: for most reasonable wage growth figures (note the model-impled one is 1.066), the scaling factor is below 2.

Figure A.E.7: Scaling factor if labour shares assumption is violated

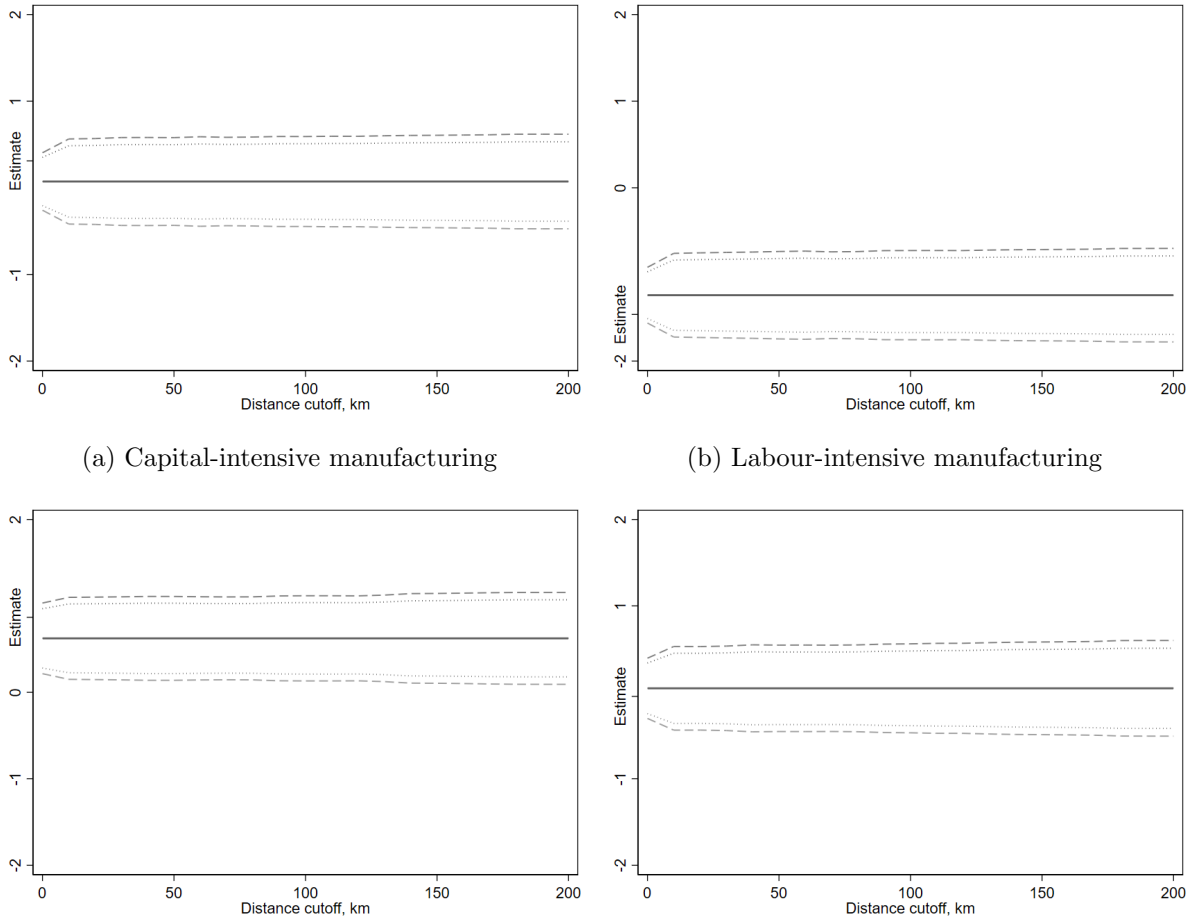


Notes: This figure shows the factor by which Ricardian differences in productivity growth are scaled up in case my assumption on equal labour shares is violated. Note that model-implied wage growth sits at 1.066.

Figure A.E.8 repeats Table 1.5.1 using Conley standard errors (Conley (1999)) and a range of different distance cut-offs. Table A.E.7 shows the long run response of sectoral productivities.

Figure A.E.9 shows output per farm in Scania, adjusting for farm size, and documents only a small and brief dip in output after the plague. Figure A.E.10 decomposes the wharves producing ships for the Swedish East India Company. Ships that were built in wharves founded after the plague in plagued regions account for the majority on all accounts.

Figure A.E.8: The Plague's Productivity Effect by Sector, using Conley Standard Errors



(a) Capital-intensive manufacturing

(b) Labour-intensive manufacturing

(c) Capital-intensive agriculture

(d) Labour-intensive agriculture

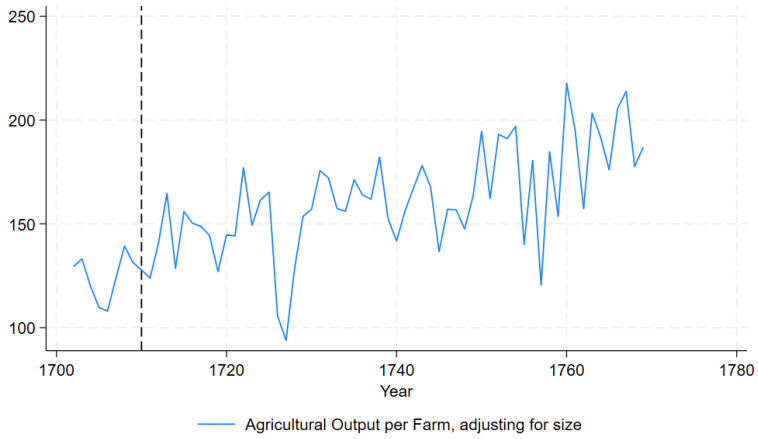
Notes: These Figures repeat columns 1, 3, 5, and 7 of Table 1.5.1 with Conley standard errors and distance cut-offs between 0 and 200 km in steps of 10 km. Distance cut-offs are denoted on the x axis and point estimates (solid line), 95% confidence intervals (dashed lines), and 90% confidence intervals (dotted lines) on the y axis. Implemented using Stata packages by Baum-Snow & Han (2024), Thiemo Fetzer, and Hsiang (2010).

Table A.E.7: Impact of plague on sectoral productivity growth in the long run

	Labour-intensive ag.	Capital-intensive ag.	Labour-intensive man.	Capital-intensive man.
	(1)	(2)	(3)	(4)
0-9 Years Post Plague	-0.553**	0.457	-0.732***	-0.226
	(0.246)	(0.292)	(0.229)	(0.252)
10-19 Years Post Plague	0.165	0.360	-1.576***	0.178
	(0.246)	(0.292)	(0.229)	(0.252)
20-29 Years Post Plague	-0.141	1.026***	-1.395***	0.298
	(0.246)	(0.292)	(0.229)	(0.252)
30-39 Years Post Plague	-0.289	0.826***	-1.697***	1.131***
	(0.246)	(0.292)	(0.229)	(0.252)
40-49 Years Post Plague	0.248	0.550*	-1.328***	1.616***
	(0.246)	(0.292)	(0.229)	(0.252)
50-59 Years Post Plague	0.647***	1.639***	-1.694***	1.773***
	(0.246)	(0.292)	(0.229)	(0.252)
60-69 Years Post Plague	1.182***	1.558***	-1.125***	1.986***
	(0.246)	(0.292)	(0.229)	(0.252)
70-79 Years Post Plague	1.287***	1.717***	-1.587***	0.629**
	(0.246)	(0.292)	(0.229)	(0.252)
80-89 Years Post Plague	1.706***	1.816***	-1.641***	0.308
	(0.246)	(0.292)	(0.229)	(0.252)
90-99 Years Post Plague	1.074***	0.912***	0.110	1.295***
	(0.246)	(0.292)	(0.229)	(0.252)
Fixed Effects:				
– Region	✓	✓	✓	✓
– Area x Year	✓	✓	✓	✓
Observations	170,403	170,430	170,430	170,430

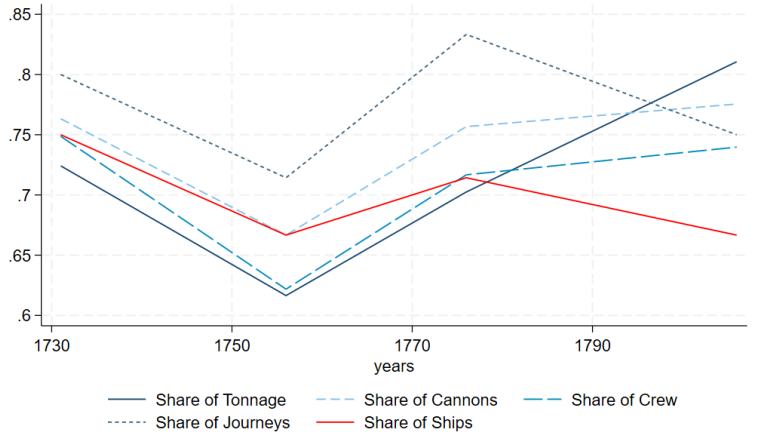
Notes: Standard errors in parentheses. * p <0.10, ** p <0.05, *** p <0.01. The dependent variable is log sectoral productivity growth. The independent variable is a plague dummy that equals 1 for plagued regions after the plague hit interacted with decade dummies after the plague. Denmark has been dropped from the sample as it reintroduced serfdom in 1733.

Figure A.E.9: Total value of agricultural production per farm, controlling for farm size



Notes: Figure A.E.9 includes all farms in Scania covered by the data and their total output.

Figure A.E.10: Share of Swedish Wharf Output by Plague Status



Notes: Based on data on the production location of all ships (n=30) operated by the Swedish East India Company. Wharves founded in plagued regions after 1714 are classified as located in plagued regions. The y axis denotes the share of tonnage, journeys, cannons, crew and ships for four points in time produced in wharves founded after the plague in plagued regions.

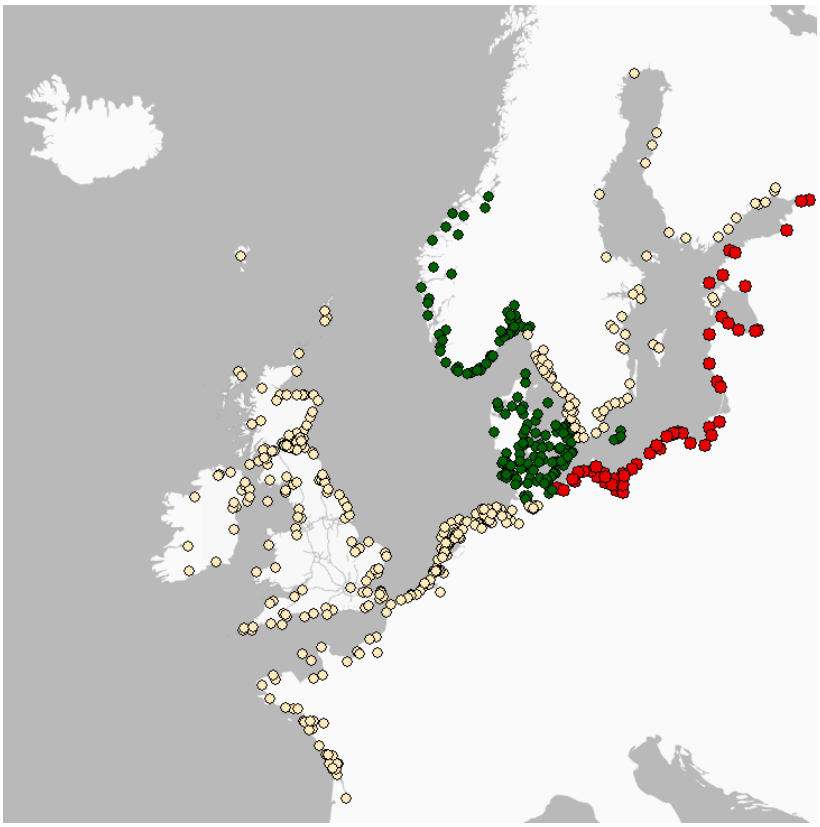
A.E.3 Evidence from Factor Frictions

Labour Frictions

There is disagreement between Raster (2023) and Peters (2022) about the presence of serfdom in Estonia, Livonia, and Pomerania. I follow Raster (2023) in coding these areas as featuring serfdom for the following reasons. Raster (2023) presents evidence for significant labour coercion in Estonia and Livonia. The German peasant-owning nobility successfully resisted Swedish attempts to abolished serfdom in 1681 (Tammisto (2020), Seppel (2019)). Laws were passed in the early 1800s that replaced serfdom with villeinage, ascribing peasants to a parish rather than an individual landowner. This, however, still left peasants without the ability to move. By 1819,

Peasant Laws had come into force, abolishing serfdom and permitting freedom of movement (Blūzma (2019)). There is evidence that even for several decades after this, mobility restrictions remained in place. (Merkel (1800)) Therefore, when judging serfdom from the perspective of mobility restrictions, Estonia and Livonia certainly were areas with serfdom. For Pomerania, Millward (1982) notes that during the 17th century, all peasants were liable for unlimited service. In Swedish Pomerania, legislation permitted the sale of landless serfs as late as the 1780s, suggesting that serfdom was present in Pomerania, too. Figure A.E.11 shows regions by their serfdom status.

Figure A.E.11: Regions by serfdom status

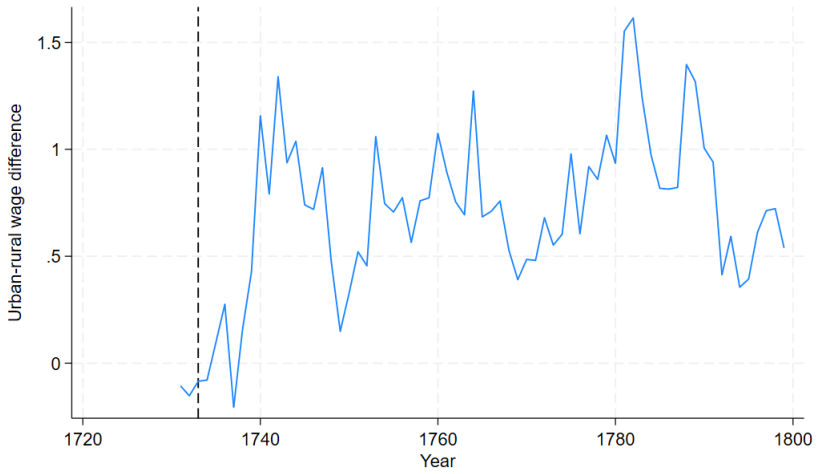


Notes: This map classifies regions by their serfdom status according to the above discussion. Yellow regions did not feature serfdom before and after the plague, whereas red regions did. Green areas re-introduced serfdom in 1733.

I use data by Gary et al. (2022) to show that there was no wage gap between unskilled urban and rural workers before the re-introduction of serfdom in Denmark in 1733, consistent with free movement between these areas. After serfdom was brought back, a wage gap appears, with scarcer urban unskilled labour compensated significantly higher. Figure A.E.12 compares nominal day wages in Danish skilling for farmhands and unskilled workers in Copenhagen. Specifically, this figure displays the ratio $wage_{ratio} = \frac{wage_{urban} - wage_{farm}}{wage_{farm}}$. While few observations

are available before 1733, the picture that emerges is one of a significant post-serfdom urban wage premium, averaging 0.7 farm day wages. With full labour mobility, peasants would have moved to Copenhagen to benefit from these higher wages. Therefore, restricted labour mobility translates into an urban-rural wage wedge. I conceptualise this by assuming a wedge of ϕ_i between urban and rural wages: $w_{iM} = (1 + \phi_i)w_{iA}$.

Figure A.E.12: Urban-rural wage differential in Copenhagen



Notes: Difference between urban and farm day wages divided by farm day wages. Data by Gary et al. (2022). The dotted line signifies the 1733 re-introduction of serfdom.

Comparing two plagued regions, one with and one without serfdom, I expect the share of workers in agriculture to be higher in the region with serfdom and the share of workers in manufacturing to be higher in the region without serfdom. Plagued cities in regions with serfdom remain relatively scarcer in labour, whereas their rural counterparts become relatively abundant in labour. The strength of the shift into capital-intensive exports therefore depends on the presence of serfdom. I revisit the result from Step #3 above for both sectors:

$$\frac{X_{ijCM}}{X_{ijLM}} = \xi w_{iM}^{-\theta(\gamma_{CM}-\gamma_{LM})} w_{iM}^{\beta_{CM}-\beta_{LM}} \frac{CMA_{jLM}}{CMA_{jCM}},$$
$$\frac{X_{ijCA}}{X_{ijLA}} = \xi w_{iA}^{-\theta(\gamma_{CA}-\gamma_{LA})} w_{iA}^{\beta_{CA}-\beta_{LA}} \frac{CMA_{jLA}}{CMA_{jCA}}.$$

With urban wages increasing more with serfdom than without, the ratio $\frac{X_{ijCM}}{X_{ijLM}}$ should increase more in regions with serfdom. Agriculture becomes relatively abundant in labour, leading to lower agricultural wage increases in areas with serfdom than in those without. Accordingly, the ratio $\frac{X_{ijCA}}{X_{ijLA}}$ should increase less in regions with serfdom. I put this prediction to the test and

estimate:

$$T_{ift} = \delta_p plague_{it} + \delta_{ps} serfdom_i \text{ x } plague_{it} + \delta_{pf} factor_f \text{ x } plague_{it} + \delta_{pfs} serfdom_i \text{ x } plague_{it} \text{ x } factor_f + \alpha_{if} + \alpha_{ft} + \gamma x_{it} + \epsilon_{ijkt},$$

where T_{ift} are exports by factor intensity, α_{if} are origin-factor fixed effects and α_{ft} are factor-time fixed effects. ϵ_{ift} denotes the error term. δ are the coefficients of interest displayed in Table A.E.8. δ_{pfs} is the most interesting coefficient as it measures how the effects of labour scarcity on capital-intensive exports vary by the presence of second serfdom. The results in Table A.E.8 confirm this expectation. The plague's effect on the capital intensity of exports is significantly higher in manufacturing under serfdom. This reflects *relative* labour scarcity.

Table A.E.8: Impact of plague on shift into capital-intensive exports by sector & second serfdom

	Manufacturing		Agriculture	
	(1)	(2)	(3)	(4)
Plague	1.232*	0.627	1.205*	0.850
	(0.707)	(0.398)	(0.681)	(0.529)
Plague x Serfdom	-2.498***	-0.930**	-2.187***	-1.325**
	(0.419)	(0.409)	(0.591)	(0.526)
Plague x Capital-Int.	-1.855***	-0.809	0.653	0.868
	(0.546)	(0.528)	(0.915)	(0.922)
Plague x Capital-Int. x Second Serfdom	2.544***	0.930*	-0.478	-0.715
	(0.539)	(0.525)	(0.849)	(0.911)
Fixed Effects:				
– Origin x Sector	✓	✓	✓	✓
– Sector x Year	✓	✓	✓	✓
Estimator	PPML	OLS	PPML	OLS
Observations	29,798	5,040	53,088	9,576

Notes: Standard errors clustered at the origin level in parentheses. * p <0.10, ** p <0.05, *** p <0.01. The dependent variable are sectoral exports in levels (columns 1 and 3) and logs (columns 2 and 4). In columns 2 and 4, values are weighted by exports in levels. The independent variables are a plague dummy, a second serfdom dummy, a capital-intensive sector dummy, and their interactions. I additionally control for growing season temperatures.

Both manufacturing sectors should see faster productivity growth under serfdom, as they are artificially labour-constrained and see larger increases in the capital-to-labour ratio. The opposite is true for agriculture, where the incentives for employing more capital are smaller under serfdom as labour is artificially abundant. Table A.E.9 tests this prediction and finds no evidence supporting this conjecture. As the shift into higher capital intensity takes place, I speculate that the missing element are learning and scale effects, hampered by insufficient population mobility.

Table A.E.9: Impact of plague on sectoral productivity growth, by second serfdom

	Labour-intensive ag.		Capital-intensive ag.	
	(1)	(2)	(3)	(4)
Plague	-0.428	0.948***	-1.169***	0.145
	(0.265)	(0.311)	(0.244)	(0.252)
Plague & Serfdom	0.822**	-0.560	-0.122	-0.125
	(0.340)	(0.399)	(0.314)	(0.323)
Fixed Effects:				
– Region	✓	✓	✓	✓
– Area x Year	✓	✓	✓	✓
Observations	41,072	41,082	41,082	41,082

Notes: Standard errors in parentheses. * p <0.10, ** p <0.05, *** p <0.01. The dependent variable is log sectoral productivity growth. The independent variable is a plague dummy that equals 1 for plagued regions after the plague hit interacted with a dummy for second serfdom.

Historical evidence confirms that agricultural productivities evolved differently under serfdom. In particular, the recovery of Russia and the Baltics shows different patterns from Sweden. Broadberry & Korchmina (2024) note that Russian agricultural output per capita was no different in 1800 than it was in 1700. There is no evidence that strip farming was abandoned in Russia, unlike in Sweden, where this change was made to save labour. Relative labour abundance in serfdom-dominated areas simply may not have necessitated changes to production methods. Crop rotation and new farming machinery also only show up towards the end of the 18th century in Russia, far later than in Sweden. Magnusson (2007) suggests that this is not driven by the availability of tools through trade, as he shows that all of Sweden could have had access to modern tools but not all areas adopted it. This suggests that for Russia, too, this was a choice. The agricultural history of Prussia, which also had serfdom, paints a similar picture of no change after the plague.

Capital Frictions

Following Antràs & Tubdenov (2025), APP can be measured as the ratio of inventories over cost of goods sold. Drawing on data for 6-digit NAICS (2017) industries in the US between 2002 and 2016, I construct this ratio from the Annual Survey of Manufacturers (U.S. Census Bureau (2002–2016)). Antràs & Tubdenov (2025) use firm-level data that I do not have access to. However, they validate their findings with industry-level controls from NBER-CES, for which ASM is the upstream source which, as opposed to NBER-CES, includes inventories, and additionally covers private firms. Antràs & Tubdenov (2025) finds that Census-based data (via NBER-CES) correlate with 0.8+ with their firm-level measure. To assign early modern goods modern industry-level APP values, I check whether a given 6-digit NAICS industry existed in 1700. 165 of 364 are assessed to have existed, clustered in food and drink, textiles, basic materials, and craft-based goods. All electrical industries, motor vehicles and aircraft, plastics and rubber, most chemicals (except botanical manufacturing and pharmaceutical preparations), petroleum products, pulp-based paper, and most modern precision manufacturing (except clocks) are dropped. I then assign goods into these industries, and find that 155 commodities are assigned into 67 industries, with the rest primary goods. My assumption is not that production technology was the same in the 1700s, but that APP contains a time-invariant component. This is reflected in the 10 goods with the highest APP, which are wines, distilled alcohols, tobacco, tiles, and some manufactured metal items. The 10 goods with the lowest APP are dominated by animal products (beef, meat, ham, bones), bread, beer, cidre, mead. Thus, APP captures a meaningful dimension of production: for distilled alcohol and wines, more time passes between input costs and sales revenue. Following Böhm-Bawerk (1889), more financially developed regions have a comparative advantage in such goods.

Table A.E.10 shows that ports with higher pre-war financial development, proxied by their export-weighted average period of production (APP), shift more strongly into capital-intensive exports after the plague, and undergo a stronger export expansion. Table A.E.11 shows that post-plague productivity growth in manufacturing correlates with pre-plague financial development.

Table A.E.10: Pre-war financial development (proxied by export-weighted APP) predicts larger treatment effects

	Sectoral Exports		Exports	
	(1)	(2)	(3)	(4)
Post Plague x APP	-0.00004	-0.00004	0.00457**	0.00463**
	(0.00004)	(0.00004)	(0.00196)	(0.00198)
Post Plague x Capital-Int. x APP	0.00015**	0.00015**		
	(0.00006)	(0.00006)		
Fixed Effects:				
– Origin x Sector	✓	✓		
– Sector x Time	✓	✓		
– Area x Year			✓	✓
– Origin x Destination			✓	✓
– Destination x Year			✓	✓
Inventory Measure	Average	End of Year	Average	End of Year
Observations	102,636	102,636	491,277	491,277

Notes: Standard errors clustered at the origin level in parentheses. * p <0.10, ** p <0.05, *** p <0.01. The dependent variable in columns 1-2 is the volume of exports by capital intensity. In columns 3-4, it is export volumes. The independent variables are a post plague dummy, the pre-1700 export-weighted APP, their interaction, and annual growing season temperatures. In columns 1-2, I also regress on the interaction of a post plague dummy with a capital-intensity dummy and that interactions interaction with pre-1700 export-weighted APP. Throughout, I also regress on the interactions of the post plague (and post plague x capital-intensity for columns 1 and 2) dummies with pre-1700 export volumes of all individual sectors.

Table A.E.11: Impact of plague on sectoral productivity growth, by financial development

	Labour-intensive ag.		Capital-intensive ag.		Labour-intensive man.		Capital-intensive man.	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Post Plague	0.05484	0.05631	0.55398**	0.55388**	-1.68205***	-1.68540***	-0.06510	-0.06578
	(0.19104)	(0.19116)	(0.22410)	(0.22425)	(0.17605)	(0.17616)	(0.18160)	(0.18171)
Post Plague x APP	-0.00001	-0.00001	0.00003	0.00003	0.00022***	0.00022***	0.00007**	0.00007**
	(0.00003)	(0.00003)	(0.00004)	(0.00004)	(0.00003)	(0.00003)	(0.00003)	(0.00003)
Fixed Effects:								
– Region	✓	✓	✓	✓	✓	✓	✓	✓
– Area x Year	✓	✓	✓	✓	✓	✓	✓	✓
Inventory Measure	Average	End of Year	Average	End of Year	Average	End of Year	Average	End of Year
Observations	48,780	48,780	48,790	48,790	48,790	48,790	48,790	48,790

Notes: Standard errors in parentheses. * p <0.10, ** p <0.05, *** p <0.01. The dependent variable is log sectoral productivity growth. The independent variable is a plague dummy, the pre-1700 export-weighted APP, and their interaction.

A.E.4 Human and Social Capital Deepening

I construct measures of goods' skill intensity and contract intensity. Following Berman et al. (1994) and Autor et al. (1998), I measure skill intensity as the share of non-production workers in total employment from the Annual Survey of Manufactures (U.S. Census Bureau (2002–2016)). I use 2002-2016 data at the NAICS 6-digit level, which covers 155 of the commodities of this paper following the mapping of goods in 1700 to modern-day U.S. industries laid out in Section 1.5.1. Following Nunn (2007) and Rauch (1999), I proxy social capital intensity with contract intensity, which I measure using the share of differentiated inputs (liberally or conservatively defined as in Rauch (1999)). I use data by Nunn (2007) at the NAICS 6-digit level. Additionally, I extend the 155 manufacturing goods, matched to NAICS 6-digit industries, by 52 raw and unprocessed goods for which I impute a value of 0 for both the shares of differentiated inputs and the share of non-production workers.

Table A.E.12 finds no shift into export goods of higher human or social capital content after the plague.

Table A.E.12: The plague leads to a shift into physical capital intensity, not into human or social capital intensity (proxied by skill and contract intensity)

	Only Manufacturing					Imputed for Raw Goods				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Post Plague	-0.89901*	-0.91600*	-0.44523	-0.44282	-0.44334	-2.47991**	-2.45044**	-1.56855**	-2.21473*	-2.15805*
	(0.54362)	(0.55347)	(0.37270)	(0.37414)	(0.37457)	(1.10691)	(1.09028)	(0.78080)	(1.21381)	(1.19053)
Post Plague x Capital-Int.	0.75102*	0.77733*	1.48996**	1.49123**	1.49197**	1.60509***	1.57660***	1.74010***	1.65517***	1.63603***
	(0.40111)	(0.40963)	(0.58206)	(0.58121)	(0.58035)	(0.55170)	(0.53425)	(0.60310)	(0.57170)	(0.55842)
Post Plague x Contract Int.	-0.00022			-0.00001		0.00015			0.00013	
	(0.00016)			(0.00007)		(0.00013)			(0.00014)	
Post Plague x Contract Int.		-0.00020			-0.00001		0.00013			0.00011
		(0.00015)			(0.00007)		(0.00012)			(0.00012)
Post Plague x Skill Int.			-4.91993**	-4.89244**	-4.89804**			-2.08948	-0.79596	-0.88498
			(1.96402)	(2.04112)	(2.04315)			(1.61679)	(1.44019)	(1.44818)
Fixed Effects:										
– Origin x Good	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
– Good x Time	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Nunn Measure	Liberal	Conservative		Liberal	Conservative	Liberal	Conservative		Liberal	Conservative
Observations	446,104	446,104	446,104	446,104	446,104	339,253	339,253	339,253	339,253	339,253

Notes: Standard errors clustered at the origin level in parentheses. * p <0.10, ** p <0.05, *** p <0.01. Results from PPML regressions, where the dependent variable is exports in levels. The independent variables are a post plague dummy, the interaction of a post plague dummy with a capital intensity dummy, and the interaction of the post plague dummy with the liberal measure (columns 1, 4, 6 and 9); or the conservative measure (columns 2, 5, 7, and 10) of the share of a good's input that are differentiated according to Rauch (1999), which in Nunn (2007) is a measure of contract intensity; and/or the interaction of the post plague dummy with the share of non-production workers averaged across 2002 to 2016, based on 6-digit NAICS industries in the Annual Survey of Manufacturers (columns 3-5 and 8-10). For columns 1-2, 4-7, and 9-10, I use Nunn's data at the 6-digit NAICS level for the 1997 classification. In columns 1-5, only the 155 goods which are covered by NAICS 6-digit industries are included, whereas in columns 6-10, I impute a value of 0 for the share of differentiated inputs and for the share of non-production workers for goods that are raw and unprocessed, bringing the total to 207 commodities.

A.E.5 Non-Homotheticity

In this Appendix, I detail the arguments made in Section 1.6 against non-homothetic demand. First, I show that my finding on productivity growth is robust to permitting non-homothetic demand. Fieler (2011) provides a trade model with non-homotheticity which allows a neat mapping onto my model under the assumption that $\theta_k = \theta \forall k$, i.e. a fixed intra-industry heterogeneity in the Fréchet distribution. Even under non-homothetic preferences, her model then breaks down to an Eaton & Kortum (2002) model:

$$X_{ijk} = \frac{\chi_k A_{ik} (w_{ik})^{-\gamma_k \theta} d_{ijk}^{-\theta}}{CMA_{jk}} X_{jk}, \tag{A.E.6}$$

where I no longer assume $X_{jk} = \alpha_k Y_j$. Note that δ_{jkt} in equation 1.4.9 still absorbs the demand side under non-homotheticity. Therefore, my findings on productivity growth are unaffected.

Next, I analyse prices, as non-homothetic demand has tractable implications for relative prices of capital-intensive goods, which should become relatively more expensive. Crucially, the productivity channel advocated for in this paper produces the opposite relative price effect. Forming the ratio of the price indices for capital-intensive and labour-intensive (agricultural) goods in equation 1.4.5, I find:

$$\frac{P_{jCA}}{P_{jLA}} = \left(\frac{\chi_{CA} \Sigma_i A_{iCA} (w_{iA})^{-\gamma_{CA} \theta} d_{ij}^{-\theta}}{\chi_{LA} \Sigma_i A_{iLA} (w_{iA})^{-\gamma_{LA} \theta} d_{ij}^{-\theta}} \right)^{-\frac{1}{\theta}}, \tag{A.E.7}$$

where I impose $w_{iCA} = w_{iLA} = w_{iA}$ due to labour mobility in the agricultural sector and $d_{ijCA} = d_{ijLA} \forall (i,j)$. This ratio takes the same form in a general model of non-homothetic demand (Fieler (2011)) as in my model. I will differentiate between three types of prices. First, the price index at the regional level, which incorporates productivities, wages, and trade costs in all exporting regions. Trade costs and productivities in unplagued exporting regions do not change, but wages may as firm market access changes given productivity growth in plagued regions. I assume these changes to be small and scaled down by trade costs to North Sea regions. Productivities in plagued exporting regions change, but their trade with Baltic importers is not observed in my data. Unobserved trade within the Baltic unequivocally lowers the relative price of capital-intensive goods via the productivity and labour cost channels. Price data at the regional level should therefore show a relative decline in the price of capital-intensive goods. Note, however, that this confounds demand- and supply-side factors.

I assemble price data at the regional level from Allen & Unger (2018).[9] For plagued Danzig and unplagued Amsterdam, their dataset holds annual prices on 77 products, which I classify according to Appendix Table A.A.3.[10] I show in Table A.E.13 that labour-intensive agricultural and manufacturing prices rise in Danzig compared to Amsterdam and capital-intensive agricultural prices fall.

Table A.E.13: Product-level prices by sector and year in Danzig and Amsterdam

	Labour-int. Agriculture				Capital-int. Agriculture				Labour-int. Manufacturing
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Post Plague	10.53***	32.82	114.2*	702.0***	-41.24***	-59.72	-16.44**	4.615***	0.665***
	(3.15e-14)	(21.90)	(16.61)	(8.00e-12)	(6.44e-14)	(10.04)	(0.663)	(1.99e-14)	(3.70e-15)
Fixed Effects:									
– City	✓	✓	✓	✓	✓	✓	✓	✓	✓
– Year		✓	✓			✓	✓		
– Product			✓				✓		
– Product x Year				✓				✓	
Observations	4140	4140	4140	452	5116	5116	5116	248	191

Notes: Standard errors clustered at the city level in parentheses. * p <0.10, ** p <0.05, *** p <0.01. The dependent variables are prices of products by sector in Amsterdam and Danzig. The independent variable is 1 after the plague in Danzig in 1709. The second, third, and fourth specifications cannot be estimated for labour-intensive manufacturing due to lack of observations. No data on capital-intensive manufacturing products are contained in the data set.

Voigtländer & Voth (2012) make the opposite prediction: a hump-shaped pattern of prices for plagued regions, with prices of capital-intensive goods increasing relative to those of labour-intensive goods. My results reject this prediction. If demand shifted relatively into capital-intensive agriculture, this should have increased its relative price. Instead, the relative price decline of capital-intensive goods supports a supply-driven explanation. These price results also reconcile the evidence that after the plague consumers received a higher share of their calories from meat and dairy (Broadberry et al. (2014)) with my assumption of homothetic demand. As capital-intensive agricultural prices drop under constant consumption shares, consumers mechanically demand higher quantities of meat and dairy. Accordingly, non-homotheticity is not required to make sense of this finding.

Owing to the exceptional granularity of the Soundtoll data, I can also analyse export and import prices, which are the other two types of prices analysed. I construct prices at the bilateral level by unifying historical measures of account and dividing duty amounts by the recorded weight (available for 46% of passages), thus creating a comparable price per kilogram. I will refer

[9]No data on goods I classify as capital-intensive manufacturing are containted in the data set and few for labour-intensive manufacturing.

[10]While the effects of labour scarcity were passed on to Amsterdam through trade, Amsterdam itself was not affected by the plague and faced higher trade costs to the regions that were plagued.

to prices of goods sent to ports as import prices and to prices of goods sent from ports as export prices.

Import prices are only indirectly affected: productivities and trade costs in North Sea regions do not change; the indirect effects of firm market access changes on wages are small and weighted down by trade costs. In Table A.E.14, I control for the supply side and show that relative prices of capital-intensive goods are unaffected by the importer's plague status. These importing prices are the relevant metric in my test for non-homotheticity below.

Table A.E.14: Good-level logarithmic import prices

	(1)	(2)	(3)	(4)
Post Plague	0.281***			
	(0.0597)			
Post Plague x Capital-Int.	-0.0997		0.0875	
	(0.116)		(0.136)	
0-10 Years Post Plague		0.143		
		(0.124)		
11-20 Years Post Plague		0.386***		
		(0.0832)		
21-30 Years Post Plague		0.493***		
		(0.0982)		
>30 Years Post Plague		0.248***		
		(0.0703)		
0-10 Years Post Plague x Cap-Int.		0.0200		0.303
		(0.224)		(0.195)
11-20 Years Post Plague x Cap-Int.		-0.0243		0.0534
		(0.129)		(0.176)
21-30 Years Post Plague x Cap-Int.		-0.122		0.0776
		(0.123)		(0.138)
>30 Years Post Plague x Cap-Int.		-0.107		0.0603
		(0.117)		(0.152)
Fixed Effects:				
– Area x Sector x Year	✓	✓	✓	✓
– Origin x Destination x Year			✓	✓
– Origin x Destination x Sector	✓	✓	✓	✓
– Origin x Sector x Year	✓	✓	✓	✓
Observations	132411	132411	89334	89334

Notes: Standard errors clustered at the destination level in parentheses. * p <0.10, ** p <0.05, *** p <0.01. The dependent variables are prices of goods by sector, origin, destination, and year as recovered from the Soundtoll data. Specifically, the recorded duty amount per good is divided by the weight whenever it is recorded (46% of observations), thus creating comparable prices per kilogram. The independent variables are a post plague dummy that is 1 after the plague and the interaction of this post plague dummy with a dummy for a capital-intensive sector. Columns 2 and 4 split up this dummy and the interaction by time horizon after the plague.

Export prices are directly affected by both the productivity and labour channels. If wages rise in plagued ports i and $\gamma_k < \gamma_{k'}$, the ratio $\frac{P_{jk}}{P_{jk'}}$ falls. Even if $\gamma_k = \gamma_{k'}$, the productivity channel will still produce a relative drop of capital-intensive export prices. In Table A.E.15, this prediction is confirmed: the (relative) export prices of capital-intensive goods shipped from plagued origins decline significantly. When analysing the time path of this price diversion, I find that the relative drop in prices of capital-intensive goods intensifies over time.

Table A.E.15: Good-level logarithmic export prices

	(1)	(2)	(3)	(4)
Post Plague	0.526***			
	(0.185)			
Post Plague x Capital-Int.	-0.511**		-0.664**	
	(0.224)		(0.298)	
0-10 Years Post Plague		0.253**		
		(0.103)		
11-20 Years Post Plague		0.501***		
		(0.169)		
21-30 Years Post Plague		0.546***		
		(0.183)		
>30 Years Post Plague		0.548***		
		(0.205)		
0-10 Years Post Plague x Cap-Int.		-0.246		-0.375*
		(0.153)		(0.224)
11-20 Years Post Plague x Cap-Int.		-0.558***		-0.814***
		(0.203)		(0.296)
21-30 Years Post Plague x Cap-Int.		-0.639**		-0.801**
		(0.248)		(0.341)
>30 Years Post Plague x Cap-Int.		-0.505**		-0.657*
		(0.252)		(0.340)
Fixed Effects:				
– Area x Sector x Year	✓	✓	✓	✓
– Origin x Destination x Year			✓	✓
– Origin x Destination x Sector	✓	✓	✓	✓
– Destination x Sector x Year	✓	✓	✓	✓
Observations	118108	118108	73091	73091

Notes: Standard errors clustered at the origin level in parentheses. * p <0.10, ** p <0.05, *** p <0.01. The dependent variables are prices of goods by sector, origin, destination, and year as recovered from the Soundtoll data. Specifically, the recorded duty amount per good is divided by the weight whenever it is recorded (46% of observations), thus creating comparable prices per kilogram. The independent variables are a post plague dummy that is 1 after the plague and the interaction of this post plague dummy with a dummy for a capital-intensive sector. Columns 2 and 4 split up this dummy and the interaction by time horizon after the plague.

Comparing import prices and export prices allows me to disentangle supply-side from demand-side factors. The null finding for import prices supports my claim that there is no evidence for

non-homothetic demand; the significant findings on export prices support my claim of supply-side changes after the plague.

Finally, I directly test for a demand shift in plagued regions. Instead of absorbing non-homothetic demand structures or comparing implied price changes, I test for the presence of non-homothetic demand. I form the ratio of equation A.E.6 between any two sectors k and k', where I need to asssume $w_{ik} = w_{ik'}$ and $\gamma_k = \gamma_{k'}$, and take the ratio of this fraction between a pre-period p, 1700-1710, and any year t after the plague:

$$\frac{X_{ijkt}X_{ijk'p}}{X_{ijk't}X_{ijkp}} = \frac{A_{ikt}A_{ik'p}}{A_{ik't}A_{ikp}}\frac{X_{jkt}X_{jk'p}}{X_{jk't}X_{jkp}}\frac{CMA_{jk't}CMA_{jkp}}{CMA_{jkt}CMA_{jk'p}}. \tag{A.E.8}$$

Relative export volumes evolve as a function of relative productivity changes in the exporting region and two demand-side factors: relative demand and relative market access. In my analysis, I form the left-hand side based on trade data and absorb relative productivity changes with exporter-time fixed effects. The argument made above about the behaviour of relative import prices under non-homotheticity simplifies the demand side further. As consumer market access is a price index, the finding that for equal labour shares and wages in a region relative (import) prices do not change carries over to the ratio of consumer market access. This is true for import prices coming from unplagued North Sea ports, which is sufficient here as that is the only trade I observe. Thus, equation A.E.8 simplifies as the final fraction drops out. In this case, I can directly test for the presence of non-homothetic demand while controlling for the supply side.

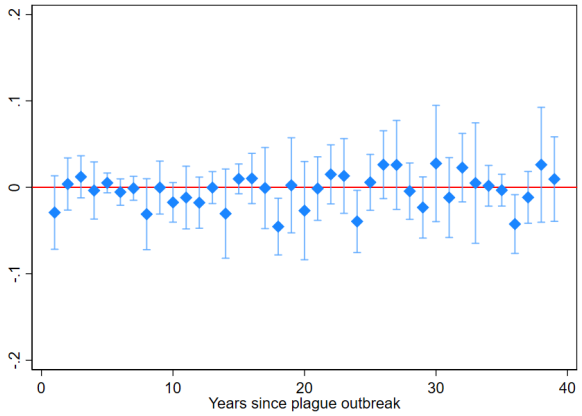
While the assumption on constant relative prices simplifies the demand side of equation A.E.8, it actually works against me in that any changes in relative demand will be seen as evidence for non-homotheticity, even if they arise from relative market access changes. The only possibility for this ratio to obscure underlying relative demand shifts is a change in relative demand that is offset by a change in relative consumer market access.

I define the baseline period p as the years 1700 to 1710 and let t be any year after 1710. In essence, the variable I regress on is a double ratio: relative exports to j compared between any post-plague year t and the pre-plague period. To capture changes in the import ratio indicative of non-homotheticity, I present event study results in Figure A.E.13 where I regress on the log of the left-hand side of equation A.E.8. The independent variable is a dummy for a plague outbreak in the importing region.[11] I absorb the productivity ratio by origin-time fixed effects and the time-invariant component in relative demand by destination fixed effects. I find no evidence of

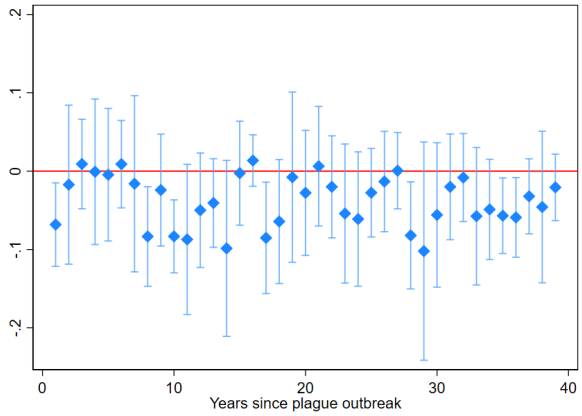
[11] I also check that there is no evidence for non-homotheticity when using the continuous plague treatment derived in Section A.C.5. Results available on request.

plague-related changes in relative demand.

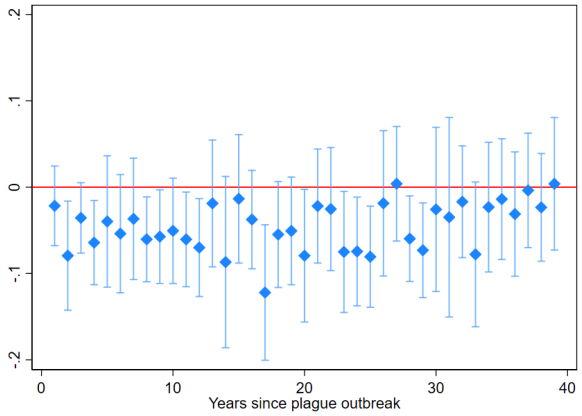
Figure A.E.13: Event studies testing for non-homotheticity



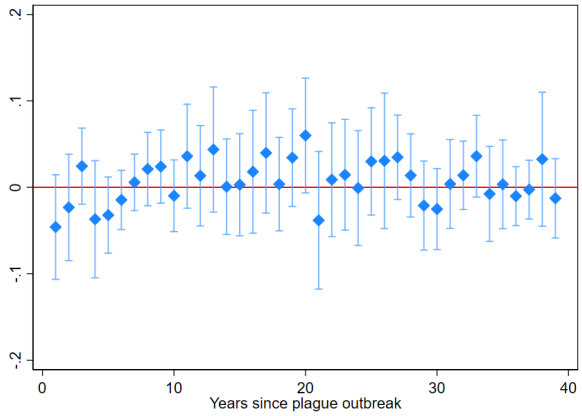
(a) Capital-int. & Labour-int. Manufacturing



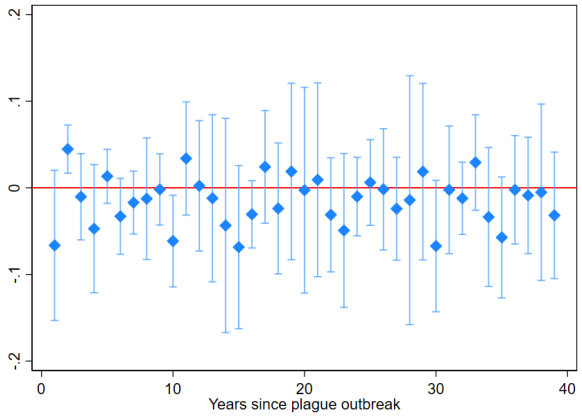
(b) Capital-int. Manufacturing & Labour-int. Agriculture



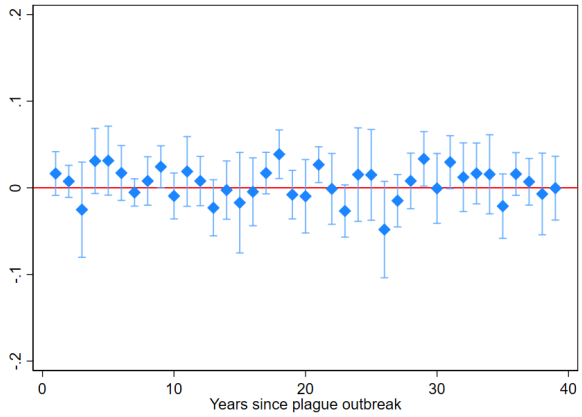
(c) Capital-int. Manufacturing & Agriculture



(d) Capital-int. Agriculture & Labour-int. Manufacturing



(e) Capital-int. Agriculture & Labour-int. Agriculture



(f) Labour-int. Manufacturing & Agriculture

Notes: Event study on the log of the left-hand side of equation A.E.8. The independent variable is a plague dummy on the importer side. Origin-time and destination fixed effects are included and standard errors are clustered at the destination level.

Appendix A.F Counterfactuals

This appendix provides details on the counterfactuals provided in Section 1.7. Constructing export shares, the outcome of equation 1.3.2, according to gravity equation 1.4.8, I find that shares can be expressed as:

$$s_{ijk} = \frac{X_{ijk}}{\Sigma_i X_{ijk}} = \frac{A_{ik}(w_{ik})^{-\gamma_k\theta} d_{ijk}^{-\theta}}{\Sigma_i A_{ik}(w_{ik})^{-\gamma_k\theta} d_{ijk}^{-\theta}}. \tag{A.F.1}$$

When plugging in equation 1.4.7 for wages and equations 1.4.5 for firm and consumer market access, market shares can be written as a function of productivities, sectoral employment, trade costs, wages, and total income.

To shut down the productivity channel, I assume A_{ik} to be fixed. I also assume fixed trade costs d_{ijk} and destination income Y_j. Only trade between the Baltic and North Seas is observed, and the plague struck almost exclusively on the Baltic Sea, rendering this a reasonable assumption.

In the estimation equation, destination-sector-time fixed effects absorb changes in the denominator. When shutting down productivity growth and ignoring the variation that will be absorbed by fixed effects, one can write $s^c_{ijkt} = \left(\frac{L_{ikt}}{L_{ikPRE}}\right)^{\frac{\gamma_k\theta}{1+\gamma_k\theta}} s_{ijkPRE}$.

I model the population recovery from L_{iPOST} to L_{it} to happen exponentially over 40 years. This is the upper bound, as by 1750 plagued cities had recovered their populations. I assume that after 40 years, regional and not just urban populations had recovered. This yields:

$$s^c_{ijkt} = (1 - m_i)^{\frac{\gamma_k\theta}{1+\gamma_k\theta}\left(1-\frac{t}{40}\right)} s_{ijkPRE}.$$

Using these counterfactual destination market shares, one can follow the multiplicative structure of this model's gravity equation to determine the average productivity growth required to bridge from the counterfactual export contraction to the observed export expansion. The average point estimate for counterfactual export shares in the event study is a market share contraction of 1.03 p.p.; the observed average market expansion is 0.98 p.p. Average productivity needs to rise by about 210% to arrive at this export expansion. This - admittedly high - estimate mirrors the magnitudes found in the TFP estimation in Table 1.5.2, where the point estimate for labour-(capital-)intensive TFP growth (column 1) implies TFP growth of 105% (205%).

Appendix B

Appendix for Chapter 2

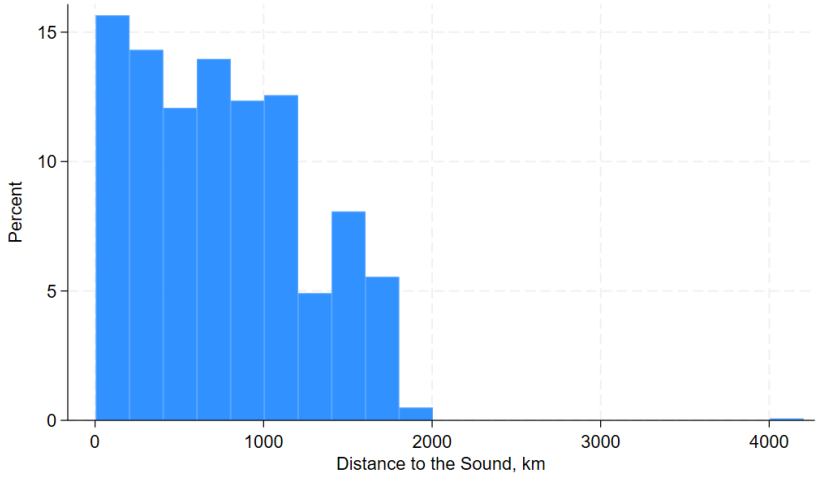
Appendix B.A Data Appendix

In this Appendix, we provide more detail about our data set construction (Section 2.3 above). Regarding the possible circumvention of the Sound, the main alternative to passing the Sound was sailing through the Great Belt. Degn (2017) makes a number of arguments for why this does not pose a major threat to our data set. Precisely because the Great Belt made it possible to avoid paying the Sound Toll duties, the Danish king introduced a prohibition on sailing through the Belt, directed specifically against the Prussian towns, extending this prohibition later on to all foreign nations. The author concludes that skippers only reluctantly evaded the Sound; it was particularly difficult for foreigners to evade the Sound. Finally, an inspection of toll records from the Great Belt reveals that the number of passages is small compared to the number of voyages through the Sound. Gøbel (2010) summarises that 85% of traffic through the Great Belt were small Danish coastal ships on short journeys. Other straits, such as the Little Belt between Jutland and Funen, were difficult and dangerous to navigate for larger ships such as the ones used for trading with the West and transport over land was expensive. According to Bergsøe (1853), the toll revenues from the Great and Little Belts combined equalled only a few percent of the Sound Toll revenues. We thus conclude that the difficulty of navigating the Great and Little Belt and the prohibition on doing so imply only a small number of unobserved vessels in our data. Most of those unobserved vessels likely were engaged in short-distance coastal trade, and below we show robustness to excluding any short-distance connections. Note further that trade over land faced a high number of tolls: whereas a ship from Danzig to Amsterdam needed to only cross one border, transport over land was way less attractive due to the high number of land borders to be traversed.

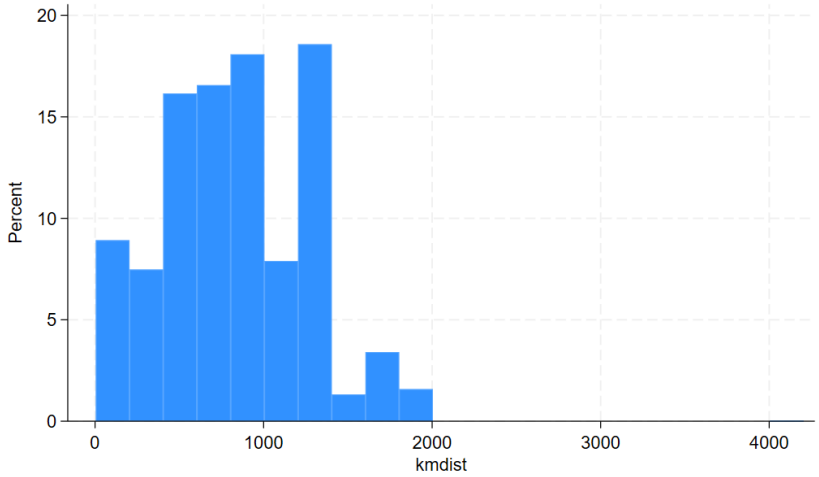
Furthermore, the Soundtoll data only record one departure and one destination. In theory, if a ship from London sails to Danzig and then onward to Saint Petersburg, it is likely that Danzig will be recorded as the destination. This could be a problem if ships tend to do stopovers just before or just after the Sound and would lead to an overrepresentation of such cities. If this were the case, it should be visible in the distribution of distances of origins and destination cities from the Sound, where we would expect short distances to be relatively overrepresented. We examine this by cutting trips from city A to city B into a trip from A to the Sound and from the Sound to B. We then assemble a data set of how often ships go on these trips. This allows us to observe how often ships sailed to different distances from the Sound. We then compare this to the "raw"(not weighted by trips) distribution of cities' distances from the Sound. If ships tended to stop just before or just after the Sound, we would expect the trip-weighted distribution to have more mass at shorter distances. As Figure B.A.1 shows, the converse is true: If anything, actual trip data show fewer short distances than one would expect given the geographic distribution of cities. There are also other arguments against the stopover issue: We see many trips with Amsterdam or Rotterdam as port of departure, a town East of the Sound as destination, and home town Bremen or Hamburg. Given that Hamburg or Bremen are much closer to the Sound than Amsterdam or Rotterdam, the latter must have been the actual origin, and not merely the last stop before the Sound. Of course, we cannot rule out multi-stop trips, e.g. a ship sailing from Hamburg to Amsterdam and then on towards the Baltic. In the data, there are few instances of several cities being recorded as the origin or destination; as these are only 412 out of 2.2 million observations, we drop these as it is impossible to ascertain which goods onboard originated in or were destined for which city. Similarly, we drop observations for which several home towns are recorded. We are inclined to interpret the occurrence of multiple recorded origins and destinations as suggesting that when a single origin or destination is recorded, this is indeed the voyage's single origin or destination.

Since the vast majority of trade flows registered at the Sound connect coastal areas in Northern and Western Europe, we imposed several geographical restrictions. We only keep cities to the North of Bayonne and drop parts of Southern France and Italy that are to the north of Bayonne, but not on the Atlantic coast, as these rarely travel far enough to cross the Sound. We further restrict our sample to cities to the west of Saint Petersburg as close to no trade comes from further inland, with the exception of Arkhangelsk, which was frequented when the Baltic harbours were inaccessible for Russia. The Northern limit we choose is the Arctic Circle and few locations are further North. To the West, our cut-off is Ireland. Only a handful of voyages

Figure B.A.1: Theoretical and observed distributions of distances to the Sound



(a) Theoretical distribution.



(b) Observed distribution.

Notes: Panel A displays the percentage of cities that fall within each distance bin of the Sound. Panel B displays observed distances to the Sound by cutting a trip from A to B into two trips, one from A to the Sound and one from the Sound to B. The distribution of these distances is shown in Panel B.

crossing the Sound originated in North America. By these selections, we cut about 20% of cities and 8% of observed trade flows. In the next step, we only keep cities that are either 25 km or less away from the coast or are 100 km or less away from the coast, but only 25 km or less away from a major river (namely Loire, Seine, Rhine, Elbe, Oder, Vistula, Neman, and Daugava). By this selection, we cut a further 40% of cities and 3% of observed trade flows. This is again motivated by the fact that shipments from these 40% of cities far inland are rare in our data. These steps drop cities that rarely show up in the data, which almost exclusively feature coastal cities in the Northern half of Europe.

Appendix B.B Further results and robustness checks

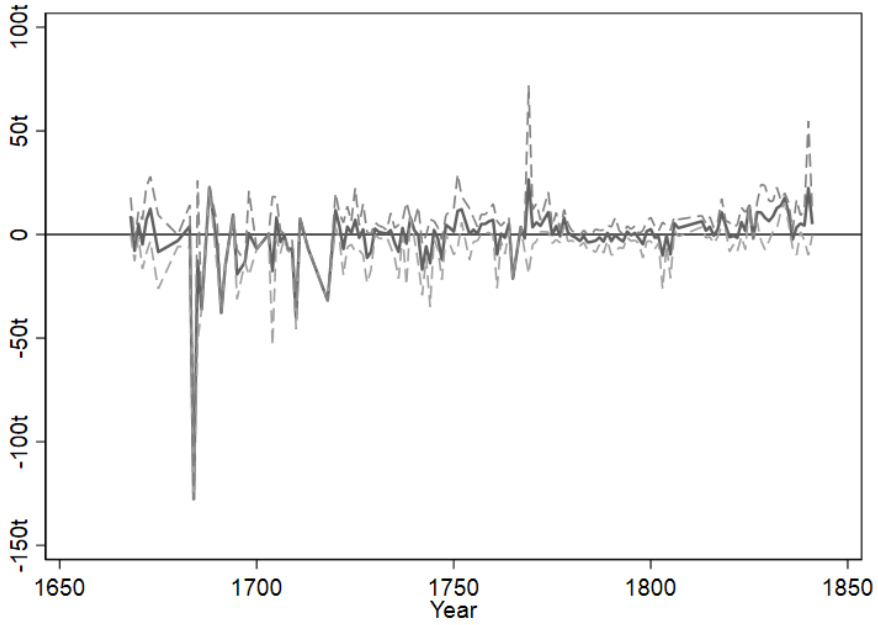
Table B.B.1 details all the Hansa (based Dollinger (1964)) and Kontor (from Hammel-Kiesow (2000)) cities in our dataset. We only list cities that satisfy our sample selection criteria outlined above.

Table B.B.1: List of Hansa and Kontor cities appearing in the trade data

City	Hanse	Kontor	City	Hanse	Kontor
Aabo	0	1	Anklam	1	0
Amsterdam	0	1	Braunsberg	1	0
Antwerpen	0	1	Bremen	1	0
Bergen	0	1	Buxtehude	1	0
Bordeaux	0	1	Danzig	1	0
Boston	0	1	Elbing	1	0
Bourgneuf	0	1	Greifswald	1	0
Brugge	0	1	Grevenstein	1	0
Colchester	0	1	Groningen	1	0
Dordrecht	0	1	Halmstad	1	0
Edinburgh	0	1	Haltern	1	0
Falsterbo	0	1	Hamburg	1	0
Great Yarmouth	0	1	Kammin	1	0
Hull	0	1	Kiel	1	0
Ipswich	0	1	Kolberg	1	0
Kalundborg	0	1	Königsberg	1	0
København	0	1	Lübeck	1	0
La Rochelle	0	1	Nyköping	1	0
Landskrona	0	1	Pernau	1	0
London	0	1	Reval	1	0
Lynn	0	1	Riga	1	0
Malmö	0	1	Rostock	1	0
Nantes	0	1	Rügenwalde	1	0
Newcastle	0	1	Stade	1	0
Næstved	0	1	Stavoren	1	0
Ribe	0	1	Stettin	1	0
Rønne	0	1	Stockholm	1	0
Sandwich	0	1	Stolp	1	0
Scarborough	0	1	Stralsund	1	0
Skanör	0	1	Treptow	1	0
Sluis	0	1	Visby	1	0
Tønsberg	0	1	Weser	1	0
Ystad	0	1	Windau	1	0
Åhus	0	1	Wismar	1	0
Kalmar	1	1			

Notes: List of Hansa cities based on Dollinger (1964). Only those cities which appear in the trade data and follow our sample restrictions are shown here. Kontor cities based on Hammel-Kiesow (2000).

Figure B.B.1: Ships from Hanseatic home towns did not differ in size from other ships



Notes: Shown are point estimates and confidence intervals, where standard errors are clustered on origin-destination pairs, from a regression of ship's tonnage on a year-interacted dummy for having a Hanseatic home town, including origin-destination-time fixed effects. Historically, most cogs, the standard Hanseatic ship, ranged in tonnage from 40 to 200 tons, with few reaching up to 1,000 tons. (Rodger (2004))

In Figure B.B.1, we present regression results investigating whether ships with Hanseatic captains differed in size from other ships. We use information on tonnage, recorded for about 17% of ships, and disregard observations in the years 1808-1812 and from 1842 onwards due to a particularly small number of observations in these years. We convert the historical units in which ship tonnage is recorded into tons and regress this normed tonnage on year-interacted dummies for having a Hanseatic home town and include a complete set of origin-destination-time fixed effects. The typical Hanseatic ship, a cog, had a tonnage of between 40 and 200 tons, with some reaching 1,000 tons. (Rodger (2004)) We find that captains from a Hanseatic home town do not have ships that differ in size from other ships.

We further use data from Bosker et al. (2013), who base their population data on Bairoch et al. (1988), to investigate whether Hanseatic cities were larger than other cities. 273 cities in our trade data have matching population data. We add information on city constitutions, the presence of a university, some geographic data, and a city's legal status from Wikipedia as controls, expanding the data set provided by Bosker et al. (2013). In Table B.B.2, we use this unbalanced sample of cities to see whether Hanseatic cities were different in terms of population than other cities in our dataset. We find that they were not significantly larger than other cities,

Table B.B.2: Hansa and city population

Variables	(1) log Population	(2) log Population
	800-1850	1600-1850
Hanse	0.231	0.241
	(0.221)	(0.269)
Elevation	0.024	0.023
	(0.014)	(0.015)
Suitability	-1.050	-1.178
	(0.260)	(0.289)
West of Sound	0.252	0.167
	(0.183)	(0.199)
Reichsstadt	0.896	1.037
	(0.223)	(0.304)
University	1.192	1.205
	(0.233)	(0.250)
Lübisch constitution	0.665	0.560
	(0.263)	(0.356)
Magdeburger constitution	0.799	0.782
	(0.307)	(0.377)
Observations	1,310	891
Number of cities	273	273

Notes: Column 1: Regressions of ln population of 273 cities and centuries 800 to 1800 and 1850. Standard errors clustered at city-level below estimates in brackets. Results obtained for pooled OLS. Additional control variables included, but not displayed: time dummies for centuries from 800 to 1800 (the constant captures 1850), the inverse hyperbolic sine of distance to the coast, the inverse hyperbolic sine of distance to the river, the inverse hyperbolic sine of ruggedness, the inverse hyperbolic sine of average annual temperature, latitude and longitude. Elevation and suitability are also transformed using the inverse hyperbolic sine. Column 2: Like column 1, but only for 1600-1850; time dummies for 1700, 1800 and 1850; the constant captures 1600. Standard errors clustered at city level are below estimates in brackets. Pooled OLS.

neither over the long run, nor in the time frame in which we observe trade data.

However, Table B.B.3 shows that Hanseatic cities imported more goods, though this effect is only significant in later decades. However, throughout Hansa towns imported a higher share of luxury goods, here defined as colonial goods and spices. Colonial goods and spices cover anise, arrack, cardamon, china, cinnamon, cloves, cocoa, coffee, dates, figs, frankicense, ivory, muscade, pepper, rum, safran, salt, silk, spices, sugar, tea, and tobacco. We interpret this as an indication of higher wealth in these cities, which is however not expressed in a larger population.[1]

In Table B.B.4 we present the results of Figure 2.8.2 in a table rather than a figure. In addition to the Hansa dummy, it also shows decade-interacted coefficients on log distance in equation 2.5.1. Figure B.B.2 plots these coefficients. We find negative estimates throughout, in

[1]Since we are interested in the shares of luxury goods, we base this analysis on a smaller dataset that only contains realized trips. That is, we look at the value (or the goods shares) of trade, conditional on trade happening.

Table B.B.3: Hanseatic Cities Imported More and More Luxury Goods

	Total Import Value	Import Share Colonial Goods	Import Share Spices
	(1)	(2)	(3)
1670s	.3231	-.0603	.1217**
	(.329)	(.0374)	(.0503)
1680s	3.08	.0952**	.1298**
	(4.956)	(.0424)	(.0508)
1690s	2.401	.057*	.2334***
	(4.204)	(.0307)	(.0607)
1700s	4.098	.0745**	.1789***
	(4.36)	(.038)	(.0558)
1710s	7.815	.0644***	.1913***
	(5.458)	(.0234)	(.0467)
1720s	9.019	.0717**	.0981***
	(5.532)	(.0305)	(.0367)
1730s	9.124	.1079***	.1307***
	(5.97)	(.0302)	(.0451)
1740s	12.47	.0699**	.1314**
	(7.69)	(.0279)	(.0545)
1750s	16.87*	.0966***	.1128**
	(8.769)	(.0269)	(.0536)
1760s	27.07**	.05***	.0857**
	(12.73)	(.0188)	(.0415)
1770s	27.27*	.0525***	.0426
	(13.97)	(.018)	(.0399)
1780s	32.83*	.1067***	.0943***
	(16.87)	(.0267)	(.0363)
1790s	17.14*	.0863***	.0406
	(8.787)	(.0239)	(.0317)
1800s	26.26**	.1648***	.0494
	(12.09)	(.0402)	(.0386)
1810s	27.3**	.2549***	.0617
	(10.85)	(.0506)	(.0514)
1820s	27.34*	.183***	.115**
	(14.4)	(.0433)	(.0517)
1830s	29.26	.191***	.1297**
	(18.15)	(.052)	(.0551)
1840s	44.63*	.1441***	.1302***
	(26.3)	(.0426)	(.0499)
1850s	33.12*	.0983***	.1064**
	(17.66)	(.0348)	(.0471)
Fixed Effects:			
– Decade	✓	✓	✓
Variable	Import Value	Import Share, Value	Import Share, Value
Goods	All Goods	Colonial Goods	Spices
Observations	6,176	6,176	6,176

Notes: OLS regressions of Hansa city dummy on total value of imports, in thousand dalers, and import shares of colonial goods and spices. Standard errors clustered on importing cities. * p <0.10, ** p <0.05, *** p <0.01. The independent variable is a dummy for whether a destination city was a member of the Hanseatic League in the past, interacted with decades. The dependent variables are the total value of imports, in thousand dalers, and the import shares of colonial goods and spices.

line with typical gravity results.

Table B.B.4: Regression results depicted in Figure 2.8.2. Point estimates for β_t and λ_t in equation 2.5.1 and regression summary statistics

Decade	Variable	Point estimate	Standard Error	p value	95% Lower Bound	95% Upper Bound
1660	β_t	-.796	.191	.0000318	-1.171	-.421
1670	β_t	-.664	.145	4.458e-06	-.947	-.38
1680	β_t	-.245	.172	.15572901	-.583	.093
1690	β_t	-.24	.33	.46614528	-.887	.406
1700	β_t	-.834	.187	8.230e-06	-1.2	-.467
1710	β_t	-1.03	.189	4.797e-08	-1.4	-.66
1720	β_t	-1.23	.168	2.372e-13	-1.559	-.901
1730	β_t	-1.122	.189	2.859e-09	-1.492	-.751
1740	β_t	-.697	.178	.00008865	-1.045	-.348
1750	β_t	-.471	.166	.00445969	-.796	-.146
1760	β_t	-.058	.16	.71604806	-.371	.255
1770	β_t	.104	.161	.51940308	-.212	.419
1780	β_t	-.054	.152	.72280237	-.352	.244
1790	β_t	-.033	.147	.8209569	-.321	.254
1800	β_t	-.219	.131	.09557047	-.477	.039
1810	β_t	-.164	.122	.17964337	-.403	.076
1820	β_t	-.164	.116	.15779246	-.391	.063
1830	β_t	-.053	.122	.66278876	-.292	.186
1840	β_t	-.041	.118	.72599018	-.272	.19
1850	β_t	-.1	.123	.41790971	-.341	.141
1660	λ_t	-4.291	.474	1.377e-19	-5.22	-3.362
1670	λ_t	-3.501	.359	1.948e-22	-4.205	-2.797
1680	λ_t	-4.098	.458	3.552e-19	-4.996	-3.201
1690	λ_t	-4.756	.487	1.477e-22	-5.71	-3.802
1700	λ_t	-5.07	.493	8.995e-25	-6.037	-4.103
1710	λ_t	-2.65	.792	.00081981	-4.202	-1.098
1720	λ_t	-3.693	.411	2.499e-19	-4.499	-2.888
1730	λ_t	-3.726	.429	3.785e-18	-4.567	-2.885
1740	λ_t	-2.502	.513	1.086e-06	-3.508	-1.496
1750	λ_t	-2.555	.394	8.951e-11	-3.327	-1.782
1760	λ_t	-2.723	.435	4.032e-10	-3.576	-1.869
1770	λ_t	-2.649	.367	5.647e-13	-3.369	-1.928
1780	λ_t	-2.774	.291	1.390e-21	-3.344	-2.204
1790	λ_t	-2.869	.307	8.241e-21	-3.47	-2.268
1800	λ_t	-2.276	.338	1.555e-11	-2.938	-1.615
1810	λ_t	-2.476	.261	2.362e-21	-2.987	-1.964
1820	λ_t	-2.136	.332	1.296e-10	-2.788	-1.485
1830	λ_t	-2.091	.29	5.912e-13	-2.66	-1.522
1840	λ_t	-2.907	.236	6.175e-35	-3.369	-2.445
1850	λ_t	-3.077	.203	1.047e-51	-3.475	-2.678
# Observations	9,712,320					
Wald chi-squared	609.7					
Pseudo R-squared	0.842					

Notes: Estimation of equation 2.5.1. Presented are point estimates, standard errors, p values, and 95% confidence interval bounds for the regression results shown in Figure 2.8.2.

Our dataset only contains trade flows that cross the Sound. To investigate whether this potentially causes selection, we conducted a simulation exercise, in which we ran 100 simulations of all possible bilateral trade flows between our 1,425 cities (the theoretical sample) and observed trade flows (those passing through the Sound). We impose a positive effect for $ODHansa_{ij}$ by setting $\beta = 1$. We then run the PPML estimation of equation 2.5.1 on the theoretical simulated sample and the observed simulated sample. Figure B.B.3 shows how similar the distributions of estimates are. Further, the mean value of β estimated in the simulated sample is 0.999986 with a standard deviation of 0.002334, which compares to a mean estimate of 1.0023 with a standard deviation of 0.00315 in the observed sub-sample. Each estimate is within the 95% confidence

Figure B.B.2: Decade-interacted distance coefficients



Notes: Estimation of equation 2.5.1, including all possible bilateral trade connections by decade. PPML estimation with (origin x time) and (destination x time) fixed effects, controlling for log distance over sea interacted with decades. Shown are point estimates of these distance coefficients accompanied by confidence intervals. Standard errors clustered on (origin, destination) pairs.

interval of the other estimate. In both cases, the simulated β coefficient on $ODHansa_{ij}$ is in the range of supported estimates and no bias is detected as arising from this sample selection.

In Figure B.B.4, we use the tax value of goods rather than the number of ships as outcome variables, obtaining very similar results as in our main specification.[2]

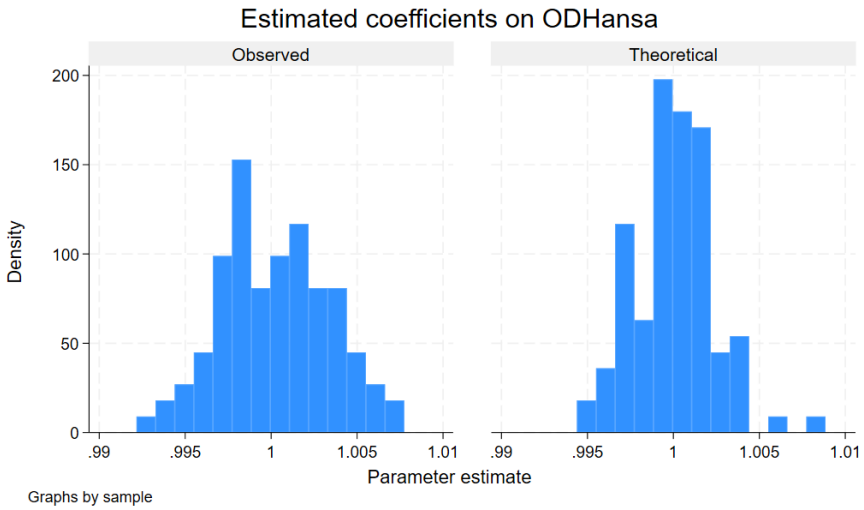
As an alternative specification to PPML, Figure B.B.5 shows coefficients estimated on simple OLS, using the log of the number of ships as outcome. Results are broadly similar, but a bit more positive towards latter decades.

To address concerns that time variation in Hansa membership matters, we restrict the pool of cities coded as Hanseatic to those attending the last Hansetag in 1669, namely Bremen, Hamburg, Lübeck, Rostock, Danzig.[3] This last general assembly is usually considered to have been the last official event of the Hanseatic League. We estimate equation 2.5.1 on this "last Hansa" dummy. The fact that the results in Figure B.B.6a are virtually identical to the ones in Figure 2.8.2 for all Hansa cities shows that time variation in Hansa membership does not appear to matter for our results. In particular, it is not the case that there was aggregate trade persistence among

[2]We bootstrap standard errors as Stata's ppmlhdfe does not provide standard errors due to separation issues.

[3]Hansetag attendance is coded from Waitz (1870). Cologne, Braunschweig, and Hildesheim also attended the last Hansetag, but are too far inland and thus not part of our sample.

Figure B.B.3: Simulation of bilateral trade: simulated Hansa effect recovered both in observed and in theoretical sample



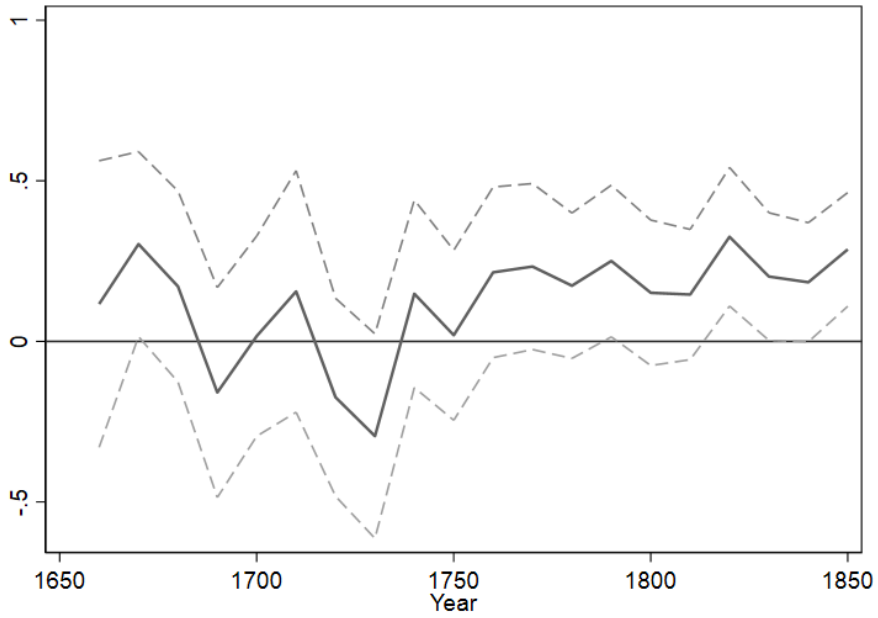
Notes: Simulation exercise, estimation of equation 2.5.1. Results from PPML regressions on simulated trade with an imposed $\beta = 1$ for the observed and the theoretical sample.

Figure B.B.4: Aggregate bilateral results on tax value



Notes: Estimation of equation 2.5.1, using tax value as the dependent variable. PPML estimation with (origin x time) and (destination x time) fixed effects, controlling for log distance over sea interacted with decades. Shown are point estimates of β_t accompanied by confidence intervals. Bootstrapped standard errors with 100 repetitions and clustered at the origin-destination pair level in parentheses. We bootstrap standard errors as Stata's ppmlhdfe does not provide standard errors due to separation issues.

Figure B.B.5: Aggregate bilateral results, OLS

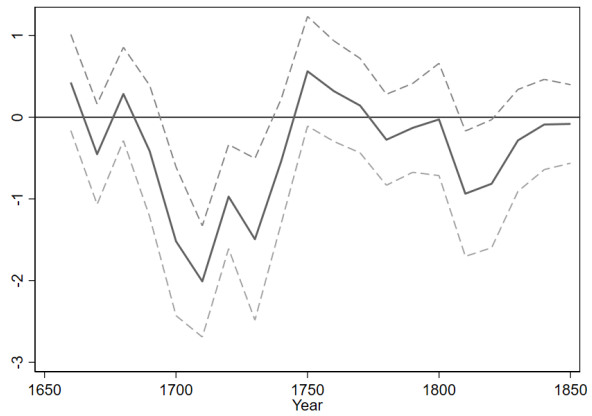


Notes: Bilateral Hansa dummy, estimation of equation 2.5.1. Results from OLS regressions on trade in logs.

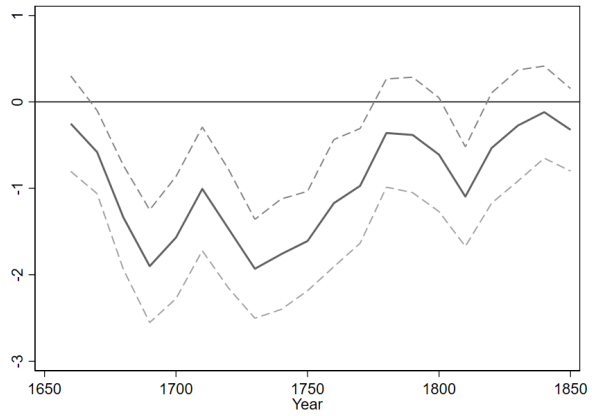
the last active members of the League. The same holds in Figure B.B.6b, where we use data from Poeck (2010) to only code cities as Hanseatic that attended any of the general assemblies between 1492 and 1516.

Table B.B.5 compares all our robustness checks on the aggregate bilateral model. We collapse the data and estimate equation 2.5.1 without the time dimension for ease of comparison. We also include results from the bias-correction in Weidner & Zylkin (2021).

Figure B.B.6: Aggregate bilateral results, only attendees of the last Hansetag and only active attendees of Hansetage



(a) Aggregate bilateral results, only attendees of the last Hansetag.



(b) Aggregate bilateral results, only active attendees of Hansetage.

Notes: Bilateral Hansa dummy, estimation of equation 2.5.1. Figure B.B.6a: Only cities attending the last Hansetag in 1669 (Bremen, Hamburg, Lübeck, Rostock, Danzig) coded as Hansa cities; Kontor cities as before. Figure B.B.6b: Hansetag attendance between 1492 and 1516 from Poeck (2010). A total of 49 cities in our sample actively attended the League's general assemblies in this period; Kontor cities as before. Estimation of equation 2.5.1 using PPML and a complete set of decade-interacted origin and destination fixed effects. Shown are point estimates of β_t accompanied by confidence intervals. Standard errors clustered on city pairs. Mean of flows: 2.3.

Table B.B.5: Overview of aggregate bilateral results without time dimension

	Number of ships				Value
	(1)	(2)	(3)	(4)	(5)
Hanseatic Route	-0.0384	-0.0384	0.173	-0.769***	0.00980
	(0.113)	(0.128)	(0.212)	(0.269)	(0.244)
Log Distance	-3.300***	-3.300***	-3.323***	-3.007***	-0.429
	(0.257)	(0.308)	(0.258)	(0.226)	(0.628)
Fixed Effects:					
– Origin	✓	✓	✓	✓	✓
– Destination	✓	✓	✓	✓	✓
Version	Bilateral sample	Weidner	Last Hanse	Active at Hansetage	Value
Observations	485,616	485,616	485,616	485,616	485,616

Notes: Estimation of equation 2.5.1 collapsed over time. Standard errors in parentheses. * p <0.10, ** p <0.05, *** p <0.01. Standard errors clustered at the origin-destination pair level. The independent variable is the number of ships in columns (1) - (4) and the value of trade in column (5). The dependent variable is a dummy for a Hanseatic route, equal to one if both cities were either Hanseatic or Kontor in columns (1)-(3) and (5), and equal to one if both cities attended the last Hansetag in 1669 in column (4). Each coefficient therefore represents the effect of both cities forming a Hanseatic route on bilateral trade. Estimation via PPML.

We next turn to robustness checks for our results that incorporate the captain's residence. Table B.B.6 reports the detailed regression results behind Figure 2.8.3a. It shows both the triple Hansa coefficients, and the coefficients on the same harbour dummies. Not surprisingly, the latter are strongly positive throughout: Captains are more likely to facilitate trade between two cities if they reside in one of them. However, our main coefficients of interest (β_t) show that over and above this, they are also more likely to facilitate trade between two Hanseatic cities, provided they themselves hail from a Hanseatic city.

We also check whether a randomly assigned placebo network produces significant results. In a simulation exercise, we randomly assign 5% of cities to a placebo Hansa, in line with the share of cities in the Hansa-Kontor group according to Table 2.8.1. Repeating this 100 times, we run our main model equation 2.6.1, including the same harbour dummies, on each simulated placebo network, with $AllHansa_{ijn}$ defined for each placebo network instead. A histogram for the estimated coefficients on $AllHansa_{ijn}$ are presented in Figure B.B.7 below. The distribution of estimates shows that the placebo network produces no significant results. Collapsed across decades for ease of comparison, the mean estimate is -1.06641 (1.04068). We conclude that the placebo Hansa network does not display effects on trade.

Figure B.B.8 shows results from using tax value rather than the number of ships to measure trade. In Figure B.B.9, we run the model with OLS, using the logged number of ships as outcome.

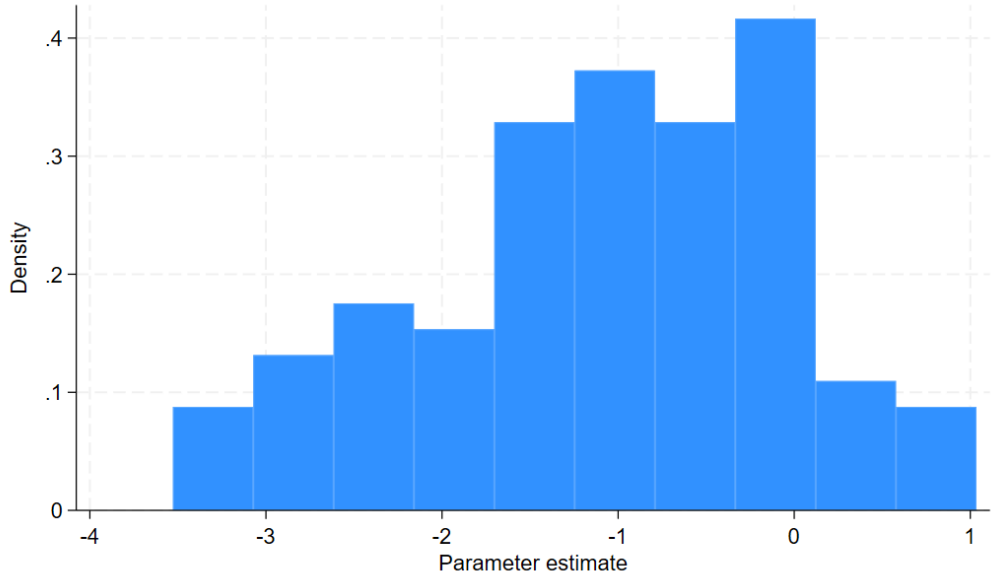
As described in the paper, the Sound Toll data only features trade flows which enter or leave the Baltic Sea through the Sound near Helsingor. Because of this, we likely miss some short-distance trade around the Sound, as land transportation and other waterways into the

Table B.B.6: Regression results depicted in Figure 2.8.3a. Point estimates for β_t and δ_t in equation 2.6.1 and regression summary statistics

Decade	Variable	Point estimate	Standard Error	p value	95% Lower Bound	95% Upper Bound
1660	β_t	.372	.27	.168	-.156	.9
1670	β_t	.274	.211	.195	-.14	.687
1680	β_t	.405	.182	.026	.048	.762
1690	β_t	.436	.152	.004	.137	.734
1700	β_t	.524	.174	.003	.184	.865
1710	β_t	.68	.28	.015	.131	1.228
1720	β_t	.693	.169	0	.361	1.024
1730	β_t	.675	.203	.001	.276	1.074
1740	β_t	.627	.17	0	.294	.96
1750	β_t	.657	.183	0	.298	1.016
1760	β_t	.533	.165	.001	.21	.856
1770	β_t	.809	.139	0	.537	1.082
1780	β_t	.462	.149	.002	.17	.753
1790	β_t	.544	.179	.002	.194	.894
1800	β_t	.362	.184	.05	.001	.723
1810	β_t	.334	.139	.016	.062	.606
1820	β_t	.518	.169	.002	.186	.849
1830	β_t	.551	.171	.001	.216	.886
1840	β_t	.548	.156	0	.242	.854
1850	β_t	.368	.093	0	.185	.55
1660	δ_t	4.464	.216	6.438e-95	4.041	4.888
1670	δ_t	3.732	.125	8.29e-196	3.487	3.977
1680	δ_t	4.366	.151	1.23e-183	4.07	4.663
1690	δ_t	4.025	.105	0	3.819	4.231
1700	δ_t	4.454	.149	3.79e-196	4.162	4.746
1710	δ_t	4.131	.13	9.22e-223	3.877	4.385
1720	δ_t	4.593	.132	1.38e-263	4.333	4.852
1730	δ_t	4.453	.157	6.04e-176	4.144	4.761
1740	δ_t	4.174	.111	0	3.957	4.391
1750	δ_t	4.279	.095	0	4.094	4.465
1760	δ_t	4.164	.08	0	4.007	4.322
1770	δ_t	4.172	.081	0	4.013	4.332
1780	δ_t	4.117	.076	0	3.968	4.266
1790	δ_t	4.021	.07	0	3.884	4.158
1800	δ_t	3.855	.065	0	3.727	3.983
1810	δ_t	3.495	.053	0	3.391	3.598
1820	δ_t	3.543	.055	0	3.434	3.651
1830	δ_t	3.34	.059	0	3.225	3.454
1840	δ_t	2.98	.057	0	2.867	3.092
1850	δ_t	2.605	.055	0	2.497	2.713
# Observations	897,043					
Wald chi-squared	10420.2					
Pseudo R-squared	0.643					

Notes: Estimation of equation 2.6.1. Presented are point estimates, standard errors, p values, and 95% confidence interval bounds for the regression results shown in Figure 2.8.3a.

Figure B.B.7: Simulation of trilateral trade: a Placebo Hansa does not produce significant results



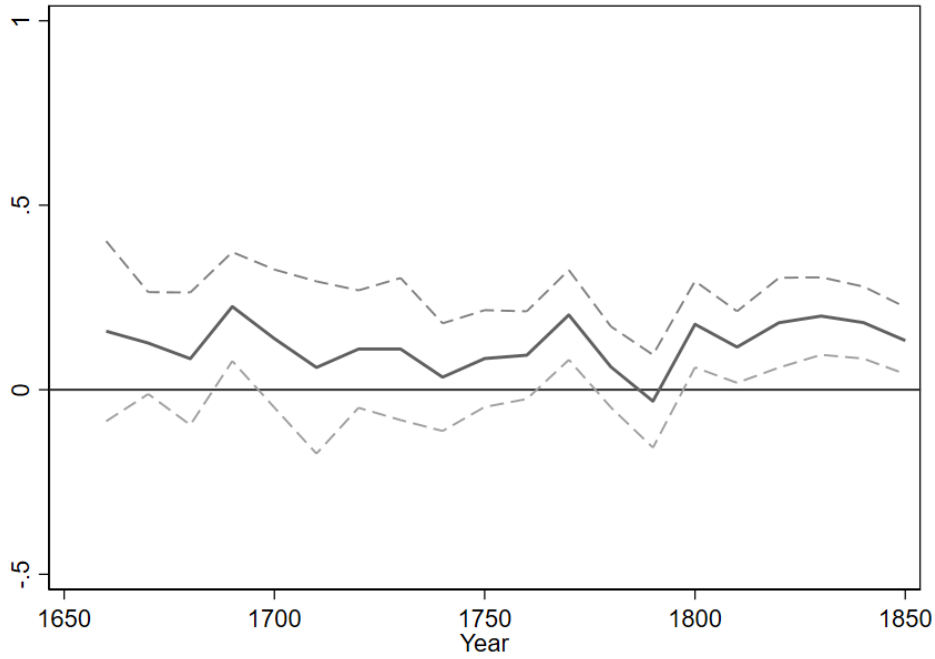
Notes: Simulation exercise, estimation of equation 2.6.1. Results from PPML regressions of observed trade on dummies indicating membership in the simulated network.

Figure B.B.8: Home town-level results on tax value



Notes: PPML estimation, including same harbour dummies (equal to one when either origin or destination coincide with the home town) of equation 2.6.1 with (origin x destination x time) and (home town x time) fixed effects, using on tax value as outcome variable. Only results for 1680 onwards are displayed, as before (1668-1679) tax value is recorded for less than 1% of observations. Shown are point estimates of β_t accompanied by confidence intervals. Standard errors clustered on city pairs.

Figure B.B.9: Home town-level results, OLS



Notes: OLS estimation of equation 2.6.1 with (origin x destination x time) and (home town x time) fixed effects in Figure B.B.17b. Results from OLS regressions on trade in logs. Shown are point estimates of β_t accompanied by confidence intervals. Standard errors clustered on city pairs.

Baltic were viable alternatives. Returning to our main specification, equation 2.6.1, Figure B.B.10 therefore shows results excluding the 10% shortest trips in order to address this issue. It is otherwise similar to previous specifications. In Figures B.B.11a and B.B.11b, we use general assembly attendance to define Hanseatic membership. Figure B.B.12 excludes the 5 largest Hansa cities. Figure B.B.13 restricts the data to frequently used connections. In all cases, our results are similar to before.

Figure B.B.14 shows our home town-level results when additionally controlling for time-interacted distances over sea between either the origin and home town or the destination and home town, whichever is shorter. Figure B.B.15 controls for a same country dummy, equal to 1 whenever origin or destination country coincide with the captain's country of residence.[4] Table B.B.7 compares all our robustness checks across models. We collapse the data and estimate equation 2.6.1 without the time dimension for ease of comparison. The mean point estimate across specifications is equivalent to a 8.8% reduction in tariffs.[5] Similarly, the mean point

[4]This is based on assuming 1871 borders as an upper bound to the nation state. We thus code the now Polish coast as German, the Baltic states and Finland as Russian, Norway as Swedish, Ireland as British, and Belgium as an independent country.

[5]Excess trade is calculated as $exp(\beta) - 1$. Following Anderson & van Wincoop (2004) and Wolf (2009), the tariff equivalent is calculated as $exp(\beta/-\sigma) - 1$. We assume $\sigma = 5$ for this exercise. The mean point estimate is 0.458. As trade costs are trilateral in these models, and as we control for a bilateral Hansa dummy throughout, this captures reduced trade costs for Hanseatic traders on Hanseatic routes.

Figure B.B.10: Home town-level results, excluding short distances



Notes: PPML estimation, including same harbour dummies (equal to one when either origin or destination coincide with the home town) of equation 2.6.1 with (origin x destination x time) and (home town x time) fixed effects, excluding the 10% of shortest trips. Shown are point estimates of β_t accompanied by confidence intervals. Standard errors clustered on city pairs.

estimate implies excess trade of 58%.

Figure B.B.11: Home town-level results, only attendees of the last Hansetag and only active attendees of Hansetage



(a) Trilateral Hansa dummy, (origin x destination x time) and (home town x time) fixed effects, only attendees of the last Hansetag.



(b) Trilateral Hansa dummy, (origin x destination x time) and (home town x time) fixed effects, only active attendees of Hansetage.

Notes: PPML estimation, including same harbour dummies (equal to one when either origin or destination coincide with the home town) of equation 2.6.1 with (origin x destination x time) and (home town x time) fixed effects, coding as Hanseatic only attendees of the last Hansetag in 1669 (Bremen, Hamburg, Lübeck, Rostock, Danzig) in Figure B.B.11a and Kontor cities as before. In Figure B.B.11b, only active attendees at Hansetage are coded as Hanseatic. Hansetag attendance between 1492 and 1516 from Poeck (2010). A total of 49 cities in our sample actively attended the League's general assemblies in this period. Kontor cities as before. Shown are point estimates of β_t accompanied by confidence intervals. Standard errors clustered on city pairs.

Figure B.B.12: Home town-level results, excluding top 5 Hansa cities



Notes: PPML estimation, including same harbour dummies (equal to one when either origin or destination coincide with the home town) of equation 2.6.1 with (origin x destination x time) and (home town x time) fixed effects, excluding the top 5 Hansa cities (Bremen, Hamburg, Lübeck, Rostock, Danzig). Shown are point estimates of β_t accompanied by confidence intervals. Standard errors clustered on city pairs.

Figure B.B.13: Home town-level results, only frequently used connections



Notes: PPML estimation, including same harbour dummies (equal to one when either origin or destination coincide with the home town) of equation 2.6.1 with (origin x destination x time) and (home town x time) fixed effects, only for frequently used connections. We keep only connections used at least 60 times throughout our roughly 180 years period, i.e. once every three years on average. Shown are point estimates of β_t accompanied by confidence intervals. Standard errors clustered on city pairs.

Figure B.B.14: Home town-level results, controlling for time-interacted distances between either the origin and home town or destination and home town (whichever is shorter)



Notes: PPML estimation including same harbour dummies (equal to one when either origin or destination coincide with the home town) of equation 2.6.1 with (origin x destination x time) and (home town x time) fixed effects. We additionally control for time-interacted distances over sea between either the origin and home town or the destination and the home town, whichever is shorter. Shown are point estimates of β_t accompanied by 95% confidence intervals. Standard errors clustered on city pairs.

Table B.B.7: Overview of home town level results without time dimension

	Number of ships					Value
	(1)	(2)	(3)	(4)	(5)	(6)
Trilateral Hansa	0.560***	0.571***	0.311***	0.502***	0.245**	0.558***
	(0.107)	(0.109)	(0.0506)	(0.135)	(0.104)	(0.197)
Same Harbour	3.563***	3.576***	N/A	3.602***	3.605***	3.816***
	(0.0408)	(0.0419)		(0.0419)	(0.0419)	(0.102)
Fixed Effects:						
– Origin x Home town						
– Destination x Home town						
– Origin x Destination	✓	✓	✓	✓	✓	✓
– Origin						
– Home town	✓	✓	✓	✓	✓	✓
Model	Bilateral x Time	Bilateral x Time	Bilateral x Time	Bilateral x Time	Bilateral x Time	Bilateral x Time
Version	Trilateral	No short trips	Lastname	Last Hansetag	Active at Hansetage	Value
Observations	103,343	98,610	452,189	103,343	103,343	97,497

Notes: Standard errors in parentheses. * p <0.10, ** p <0.05, *** p <0.01. Standard errors clustered at the origin-destination pair level. The independent variable is the number of ships or the value of trade. The dependent variable is a Hansa dummy, equal to one for trade flows on Hanseatic routes and with a Hanseatic home town or with a captain with a Hanseatic last name. Each coefficient therefore represents the effect of the Hansa network on trade. Estimation via PPML. Column 3 shows no point estimates for the same harbour dummy as the concept does not carry over to the data set on last names.

Figure B.B.15: Home town-level results, controlling for same country dummy



Notes: PPML estimation including same country dummies (equal to one when either origin or destination country coincide with the ship's country of registration) of equation 2.6.1 with (origin x destination x time) and (home town x time) fixed effects in Figure B.B.15. We choose Europe's 1871 borders as an upper bound on the nation state. Shown are point estimates of β_t accompanied by 95% confidence intervals. Standard errors clustered on city pairs.

In Table B.B.8 we address concerns that Hanseatic routes may have required particular navigational skills. If these skills had been transmitted intergenerationally among Hanseatic captains, this could drive our result that Hanseatic captains continued to trade more on Hanseatic routes after the dissolution of the League. To the contrary, Table B.B.8 finds that Hanseatic routes were not longer and connected geographically more similar cities when compared to other routes. Further, the most narrow and difficult point of any route was likely the Sound, which all routes had to pass. Finally, we show in Figure 2.8.1d that non-Hanseatic captains frequented Hanseatic routes throughout. It is implausible that these captains would not have accumulated the required navigational skills over the course of two centuries.

Table B.B.10 investigates whether our result can be explained by Hanseatic traders using more harmonised units of account on their key routes when the captain was Hanseatic as well. In essence, we estimate an equation identical to equation 2.6.1 on a data set of the same dimension, but use as the outcome measures of variation in units. We clean the information on units to arrive at 59 historical units of account, some of which are weights, others volumes and count measures. As goods may vary in the number of applicable units, we also clean the information on goods and code 227 commodities. We then construct a dataset holding two variables capturing the

Table B.B.8: Hansa routes connect geographically more similar cities

	Distance over sea	Δ Ruggedness	Δ Temperature	Δ Distance to Coast
	(1)	(2)	(3)	(4)
Hansa Route	.002	-2464***	-.1805***	-.0087***
	(.0091)	(249.8)	(.0101)	(.001)
Fixed Effects:				
– Origin	✓	✓	✓	✓
– Destination	✓	✓	✓	✓
Observations	485,616	485,616	485,616	485,616

Notes: Regressions of a Hansa route dummy on route characteristics. Standard errors clustered on origin-destination pairs. * p <0.10, ** p <0.05, *** p <0.01. The independent variable is a dummy for a Hanseatic route. The first dependent variable is distance over sea. The following dependent variables are the absolute values of the differences between origin and destination in ruggedness, average temperature, and distance to the coast.

Table B.B.9: Overview of home town level results without time dimension, adding origin x home town and destination x home town fixed effects

	Number of ships		
	(1)	(2)	(3)
Trilateral Hansa	0.150***	0.458***	0.468***
	(0.0551)	(0.124)	(0.140)
Fixed Effects:			
– Origin x Home town	✓	✓	
– Destination x Home town	✓		✓
– Origin x Destination	✓	✓	✓
– Home town	✓	✓	✓
Observations	67,136	86,335	84,598

Notes: Standard errors in parentheses. * p <0.10, ** p <0.05, *** p <0.01. Standard errors clustered at the origin-destination pair level. The independent variable is the number of ships. The dependent variable is a Hansa dummy, equal to one for trade flows on Hanseatic routes and with a Hanseatic home town. Each coefficient therefore represents the effect of the Hansa network on trade. Estimation via PPML. Number of observations varies between columns due to singletons.

variety of units used. First, the standard deviation of a unit's share by origin-destination-home town-decade observation. When on a route in a year on a ship with a home town all goods are of the same unit, this measure equals 0. The second variable counts how many units are used by origin-destination-home town-decade observation. We then regress the standard deviation of units' shares and the number of units used on a dummy for a Hansa captain on a Hansa route, in the latter case additionally controlling for the total number of ships per observed origin-destination-home town-decade observation. In both cases, we also control for the standard deviation of goods' shares or the number of different goods shipped and for the same harbour dummies capturing whether a captain's home town coincides with the origin or destination of the trip. In Table B.B.10, we find no cohesive evidence that Hanseatic captains recorded more standardised units.

Table B.B.10: No evidence that Hanseatic captains used more standardised units

	Standard Deviation of Units' Shares	# Units
	(1)	(2)
Hansa Trader on Hansa Route	.1125***	.0376
	(.0359)	(.0495)
Same Harbour	1.521***	1.255***
	(.0196)	(.0425)
Fixed Effects:		
– Origin x Destination x Decade	✓	✓
– Home town x Decade	✓	✓
Control for # Ships		✓
Control for Goods Variation	✓	✓
Observations	605,653	893,600

Notes: PPML regressions of a Hansa captain on Hansa route dummy on usage of units, controlling for same harbour dummy and goods variation. Standard errors clustered on origin-destination pairs. * p <0.10, ** p <0.05, *** p <0.01. The independent variables are a dummy for whether trade is among Hansa and Kontor cities and served by a Hanseatic captain and a dummy for a Hanseatic route, i.e. between Hansa and Kontor cities. The first dependent variable is the standard deviation of units' shares per origin-destination-home town-decade observation. If in a decade on a route all trade on a ship with a home town are of the same unit, this equals 0. The second dependent variable is the number of units used per origin-destination-home town-decade observation, where we additionally control for the total number of ships in this observation. Number of observations varies between columns due to singletons.

Similarly, we analyse commodity concentration by estimating equation 2.6.1 on the standard deviation of goods' shares and the number of different goods. In Table B.B.11, we find that Hanseatic traders traded a larger variety of goods on Hanseatic routes with Hanseatic captains.[6]

In Figure B.B.16, we investigate whether our results require both Hanseatic origins and destinations. For these results, we include dummies for whether only origin and home town or only destination and home town are Hanseatic (origin and destination Hanseatic is captured by

[6]In results available upon request, we find significant positive estimates in each decade, which speaks against decaying monopoly power as an alternative explanation for our result.

Table B.B.11: Hanseats traded a larger variety of goods on Hanseatic ships

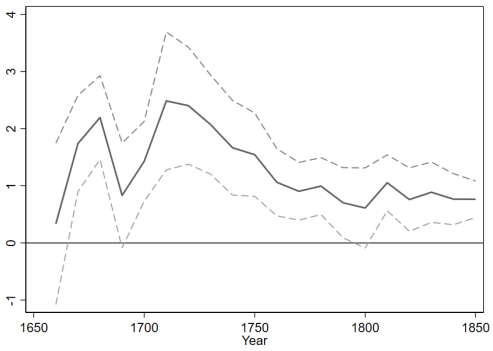
	Standard Deviation of Goods' Shares	# Goods
	(1)	(2)
Hansa Trader on Hansa Route	.0861**	.2395***
	(.0428)	(.0518)
Same Harbour	2.134***	2.352***
	(.0265)	(.0284)
Fixed Effects:		
– Origin x Destination x Decade	✓	✓
– Home Town x Decade	✓	✓
Control for # Ships		✓
Observations	595,798	893,600

Notes: PPML regressions of a Hansa captain on Hansa route dummy on variety of goods, controlling for same harbour dummy. Standard errors clustered on origin-destination pairs. * p <0.10, ** p <0.05, *** p <0.01. The independent variables are a dummy for whether trade is among Hansa and Kontor cities and served by a Hanseatic captain and a dummy for a Hanseatic route, i.e. between Hansa and Kontor cities. The first dependent variable is the standard deviation of goods' shares per origin-destination-home town-decade observation. If in a decade on a route on a ship with a home town only one kind of good is traded, this equals 0. The second dependent variable is the number of goods shipped per origin-destination-home town-decade observation, where we additionally control for the total number of ships in this observation. Number of observations varies between columns due to singletons.

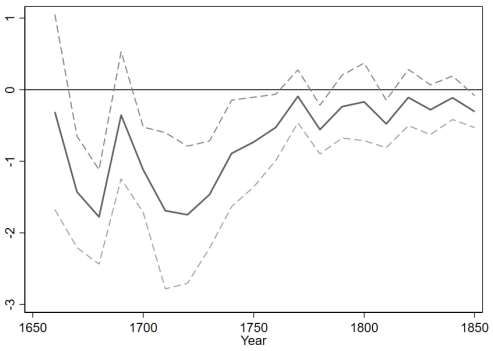
the fixed effects we have). We find strong evidence that indeed all 3 of origin, destination, and captains need to be Hanseatic: Our main variable of interest shows the same pattern as before, but the two additional dummies are mostly insignificant and sometimes even negative. This suggests that Hanseatic trading capital was only accumulated by Hanseatic captains on trips between Hansa and Kontor cities. We view this as consistent with the network effect: Capital was accumulated by captains belonging to the network trading on routes connecting the network.

In Figures B.B.17a and B.B.17b, we have split the triple Hansa dummy into a strict one (where origin, destination, and home town were all Hansa towns) and a lenient one for a pair of Kontor cities or a pair of one Hansa and one Kontor city, and a Hanseatic captain. Figure B.B.17a shows that three decades after the end of the Hanseatic League, the strict triple Hansa trading capital had fully depreciated. Figure B.B.17b shows results for the remaining triple dummy, i.e. Kontor-Kontor, Hanse-Kontor, and Kontor-Hanse trade on ships with Hanseatic captains. This decomposition suggests that there is some attenuation in our results, but that Hansa captains crowded into Hansa-Kontor trade. This results could imply that trading capital is particularly important when managing at distance.

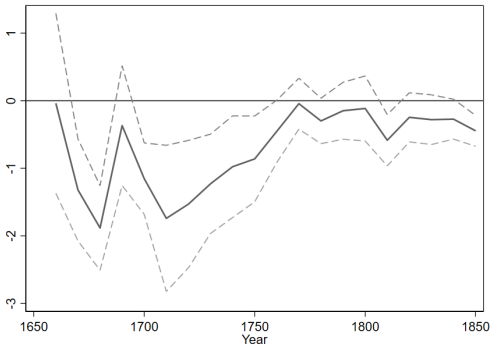
Figure B.B.16: Network results require Hanseatic captains and both Hanseatic origin and destination



(a) Estimates of β_t^1.



(b) Estimates of β_t^2.



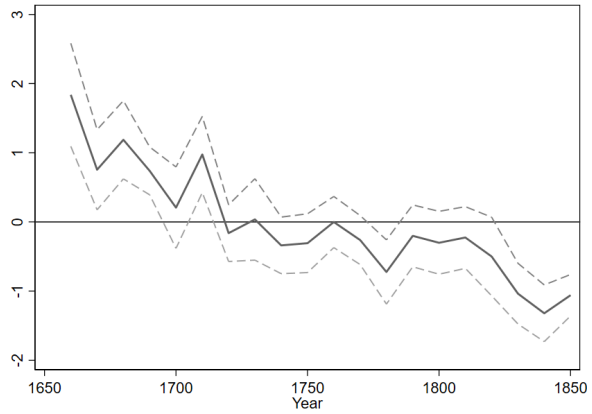
(c) Estimates of β_t^3.

Notes: For these results, we estimate:

$$T_{ijnt} = exp(\alpha + \beta_t^1 AllHansa_{ijn} + \beta_t^2 HansaOriginAndHomePort_{ijn} + \beta_t^3 HansaDestAndHomePort_{ijn} + \delta_t Harbour_{ijn} + \chi_{ijt} + \xi_{nt}) \times \epsilon_{ijnt}, \tag{B.B.2}$$

thus investigating whether the results we find require both Hanseatic origins and destinations. We find strong evidence that this is the case.

Figure B.B.17: Home town-level results, distinguishing strict Hansa dummy from Hansa-Kontor dummy



(a) Strict trilateral Hansa dummy, (origin x destination x time) and (home town x time) fixed effects.



(b) Trilateral Hansa-Kontor dummy, (origin x destination x time) and (home town x time) fixed effects.

Notes: PPML estimation, including same harbour dummies (equal to one when either origin or destination coincide with the home town) of equation 2.6.1 with (origin x destination x time) and (home town x time) fixed effects. The strict Hansa dummy is 1 whenever origin, destination, and the home town were Hansa towns. The Hansa-Kontor dummy is 1 on Hansa-Kontor, Kontor-Hansa, and Kontor-Kontor trade on ships with a Hanseatic captain. We thus find attenuation of our results for strict Hanseatic trade after three decades. Hanseatic captains crowded into Hansa-Kontor, Kontor-Hansa, and Kontor-Kontor trade. Shown are point estimates of β_t accompanied by confidence intervals. Standard errors clustered on city pairs.

Appendix B.C Home town-level Estimation- Further Results and Robustness checks

We now show that our results also hold when using specifications from existing gravity frameworks. A recent literature on multinational production, started by Arkolakis et al. (2018), models trade flows among countries and within global firms. Head & Mayer (2019) have developed an alternative specification that is also grounded in the multinational production literature and is a generalised version of the model by Arkolakis et al. (2018) which only has data for US firms. In line with their model, we estimate with origin x time and destination x home town x time fixed effects:

$$T_{ijnt} = exp(\alpha + \beta_t AllHansa_{ijn} + \gamma_t ODHansa_{ij} + \delta_t Harbour_{ijn} + \lambda_t dis_{ij} + \eta_{it} + \psi_{jnt}) \times \epsilon_{ijnt}, \tag{B.C.1}$$

where T_{ijnt} denotes the number of ships sailing from city i to city j on ships with home town city n at time t, $ODHansa_{ij}$ is a dummy for a Hanseatic route, $Harbour_{ijn}$ is a dummy equal to 1 whenever a ship's home town coincides with a journey's origin or destination, dis_{ij} is the log distance over sea between the two harbours, and η_{it} and ψ_{jnt} are origin x decade and destination x home town x decade fixed effects. ϵ_{ijnt} is the error term. The coefficients β_t on $AllHansa_{ijn}$, a dummy that is one when all of origin, destination and home town city are Hanseatic, are what we are most interested in, as it measures by how much trade on Hanseatic routes overshoots gravity predictions when executed by Hanseatic traders. Given the inclusion of the double Hansa dummy $ODHansa_{ij}$, it estimates whether Hanseatic cities continued to trade with each other via Hanseatic captains only. As before, $AllHansa_{ijn}$, $ODHansa_{ij}$, $Harbour_{ijn}$, and dis_{ij} are interacted with decade dummies.

In equation B.C.1, η_{it} accounts for the fact that Figure 2.8.1a below shows that Hansa towns' exports grow more slowly. ψ_{jnt} absorbs all destination x home town variation, which addresses Figure 2.8.1c below showing trade with the West generally grows faster. ψ_{jnt} nests a home town x decade fixed effect and thus addresses the declining Hanseatic role in shipping, seen in Figure 2.8.1b. Results are shown in Figure B.C.1a.

Specification B.C.1 is based on Head & Mayer (2019), who generalise the model by Arkolakis et al. (2018). An alternative specification is motivated by the literature on gravity with disaggregated data such as products or sectors. Based on Anderson & van Wincoop (2004), Yotov et al. (2016) suggest to model sector-specific multilateral resistance terms as exporter x product

x time and importer x product x time fixed effects. A similar specification has also been used by Arkolakis et al. (2018). The equivalent in our data are origin x home town x time and destination x home town x time fixed effects. The same harbour dummy $Harbour_{ijn}$, indicating whenever a ship's home town equals the origin or destination of a journey, is here absorbed by fixed effects. The same is true of the distances between the home town and the origin and the home town and the destination, here absorbed by μ_{int} and ν_{jnt}. We thus estimate:

$$T_{ijnt} = exp(\alpha + \beta_t AllHansa_{ijn} + \gamma_t ODHansa_{ij} + \lambda_t dis_{ij} + \mu_{int} + \nu_{jnt}) \times \epsilon_{ijnt}, \qquad \text{(B.C.2)}$$

The results from this specification are shown in Figure B.C.1b. As in our main specification in Figure 2.8.3a, we find substantial and long-lasting persistence.

Figure B.C.1: Home town-level results from multinational production and sectoral gravity models



(a) Trilateral Hansa dummy, multinational production. Estimating equation B.C.1.



(b) Trilateral Hansa dummy, sectoral gravity model. Estimating equation B.C.2.

Notes: Trilateral Hansa dummy. PPML estimation of multinational production equation B.C.1 with (origin x time) and (destination x home town x time) fixed effects in Figure B.C.1a. PPML estimation of sectoral gravity equation B.C.2 with (origin x home town x time) and (destination x home town x time) fixed effects in Figure B.C.1b. Shown are point estimates of β_t accompanied by confidence intervals. Standard errors clustered on city pairs.

Appendix C

Appendix for Chapter 3

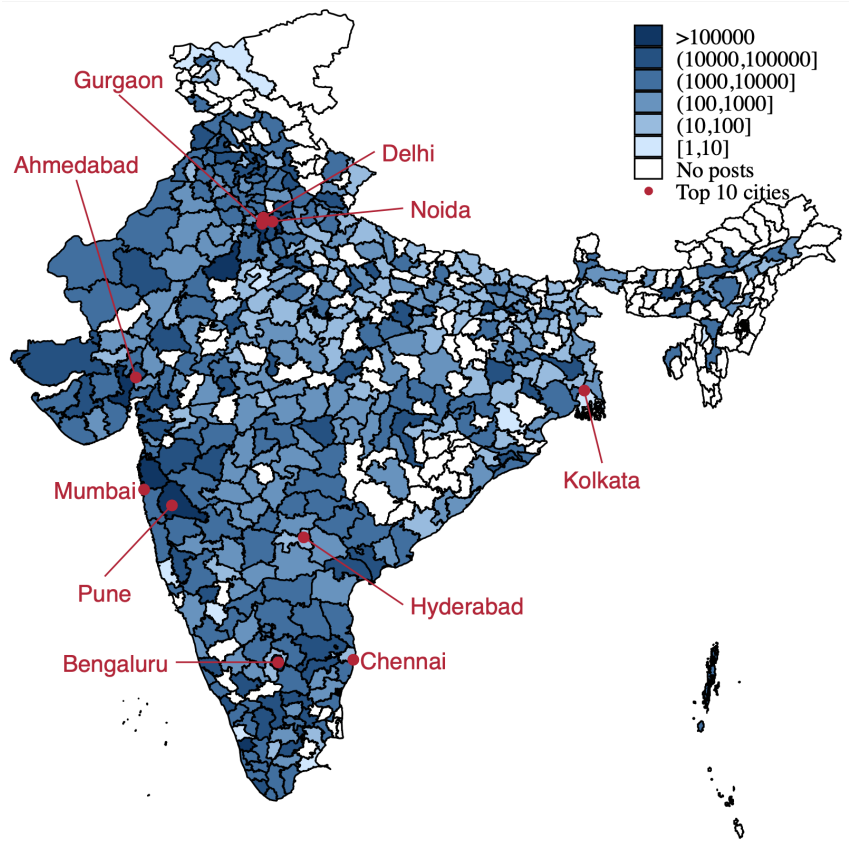
Appendix C.A Data and Descriptives

C.A.1 Construction of the vacancy dataset

The largest online job postings platform in India shared 80 percent of all job postings on its site (randomly sampled) from 2010 to 2019. All posts include text data on the job title, industry, role category, location, skills required, salary and experience ranges and educational requirements. We manually map 99 percent of role titles to the 2004 Indian National Classification of Occupations (NCO) at the four-digit level. We also manually map all industries to the 2008 Indian National Industrial Classification (NIC) at the two-digit level. We clean 95 percent of city names and add geo-locations, separating out overseas job postings. Using the geolocations, we match cities to districts, using the 2011 census. Figure C.A.1 shows the distribution of posts across districts.

Table C.A.1 summarises the number of observations in our main dataset, as well as in three other datasets that we use to gauge its representativeness (see the next section). We also use publicly-available crosswalks to translate the AI exposure measures to the Indian context. We map the 2000 Standard Occupation Classification used by Webb (2020) to the 2004 Indian National Classification of Occupations (NCO) via the 1988 International Standard Classification of Occupations (ISCO), at the four-digit level. For the Felten et al. (2018) measure, we map the 2008 ISCO to the 1988 ISCO, before again mapping onto the 2004 NCO.

Figure C.A.1: Total posts by district, 2010-2019



Notes: This map shows the distribution of our online vacancy posts across Indian districts for the entire period 2010-2019. Labels are shown for the ten cities with the largest numbers of posts.

Table C.A.1: Number of observations by data source

Online vacancy postings 2010-2019	#Firms	#Posts
Agriculture	13,811	463,675
Manufacturing	57,980	2,543,995
Services*	167,969	15,481,330
— *Financial*	*17,805*	*1,815,798*
— *Information*	*72,057*	*5,834,878*
— *Professional*	*38,533*	*834,932*
— *Other*	*106,798*	*6,995,722*
Prowess (balance sheets)	#Firms	#Observations
Agriculture	123	590
Manufacturing	2,276	11,257
Services	3,675	16,722
— *Financial*	*1,020*	*4,830*
— *Information*	*516*	*2,557*
— *Professional*	*199*	*811*
— *Other*	*1,940*	*8,524*
Surveys (demographics)	#Districts	#Households
National Sample Survey 2012	626	101,725
Periodic Labour Force Survey 2018	646	102,063

Notes: Some services firms post in multiple sub-sectors, hence the total number of services firms is less than the sum of all firms posting in the sub-sectors.

C.A.2 Representativeness of the vacancy data

In this section we evaluate the representativeness of our vacancy data in relation to the broader Indian labor market by benchmarking against widely-used administrative datasets and labor surveys. First, we consider the number of vacancies in the dataset. The Indian Labour Ministry estimates that India's services sector formally employs approximately 18.9 million workers.[1] Our unrestricted sample for 2010-19 includes approximately 15.5 million vacancies in the services sector. To compare these numbers, we need to turn the flow variable of hiring into the stock of employment. Shimer (2012) documents a job separation rate of 3.4 percent for the US. If we also assume this number for India, this implies 6.4 million job ads over 10 years. This disregards the growth of India's service sector, however, and if we assume that one quarter of job posts are replacement and three quarters are employment growth, this implies about 25.7 million job ads should be expected in the formal services sector over the course of 10 years. For this rough estimate, our coverage would be approximately 60 percent. While this back-of-the-envelope estimate is necessarily imprecise due to the lack of comprehensive worker censuses over our full sample period, we consider it nonetheless reassuring that our data are likely to reflect a substantial share of the market that we aim to understand.

We next consider the composition of the data relative to three widely used administrative datasets: firm balance sheet data from Prowess, and nationally representative labor surveys conducted in 2011-2012 (the National Sample Survey) and in 2017-2018 (the Periodic Labour Force Survey). The industry distribution of services firms in the vacancy data and the Prowess firm dataset are shown in Figure C.A.2 Panel (a). The distribution of vacancies is shown in Panel (b), alongside the distribution of equivalent white-collar services sub-industries in the pooled National Sample Survey (NSS) and Periodic Labour Force Survey (PLFS).[2] The vacancy dataset has relatively fewer finance, insurance and real estate firms than Prowess, but a greater share of vacancies in that sector relative to the representative labor surveys. The national surveys also report relatively more workers in education and transportation, likely because they include public sector workers, whereas the vacancies and Prowess balance sheet data include only private firms. Panel (c) shows the distribution of occupations in the vacancy data in contrast to the national surveys. As would be expected, the vacancy data is over-representative of high-skill white-collar jobs and under-representative of lower-skilled jobs, such as shop assistants or security

[1] See the January 2022 Quarterly Employment Survey, available at https://static.pib.gov.in/WriteReadData/specificdocs/documents/2022/jan/doc20221104101.pdf

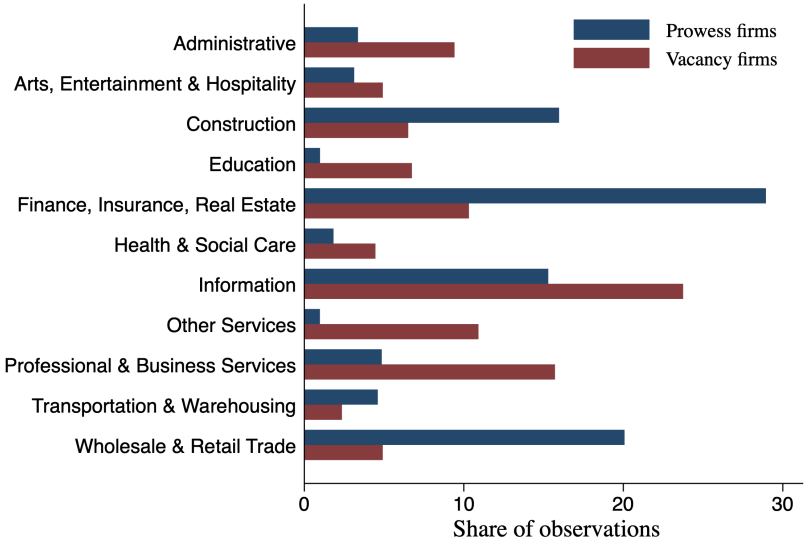
[2] We define white-collar services workers in the NSS context as salaried workers in divisions 1-5 of the 2004 Indian National Classification of Occupations, i.e. excluding agricultural, fishery, craft, manufacturing, elementary and unclassified workers.

guards, which are more typically filled through referrals and offline hiring. Panel (d) compares the number of firms over time in Prowess and the vacancy data. Prowess is not a firm census but is instead based on firms' published financial reports; it significantly over-represents India's largest firms (Goldberg et al. 2009).[3] In contrast, the barrier to appearing in our vacancy dataset is low, since the process for setting up an account and posting an online job advert is simple and does not require accounting expertise or other specialized training. We therefore see substantially more firms in the vacancy data than in Prowess, albeit with the total number of firms following a similar trajectory over time.

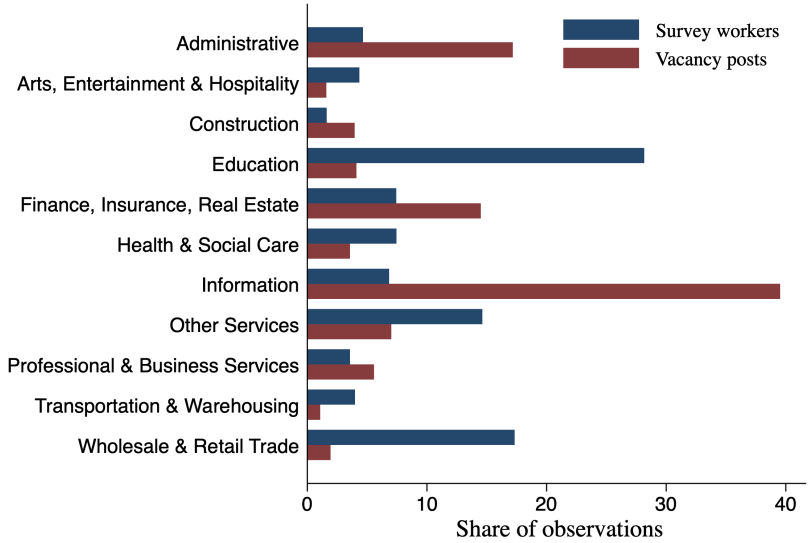
[3]Specifically, Prowess contains longitudinal balance sheet data on all publicly listed and many large private Indian firms, as long as they submit financial statements to the government. About a third of Prowess firms are publicly listed, and even among those, some do not submit financial statements to the government (Alfaro & Chari 2009). There are many firms that both are not listed and do not submit financial statements, so do not appear in the dataset.

Figure C.A.2: Comparison of vacancy data with Prowess firm-level data and labor force surveys

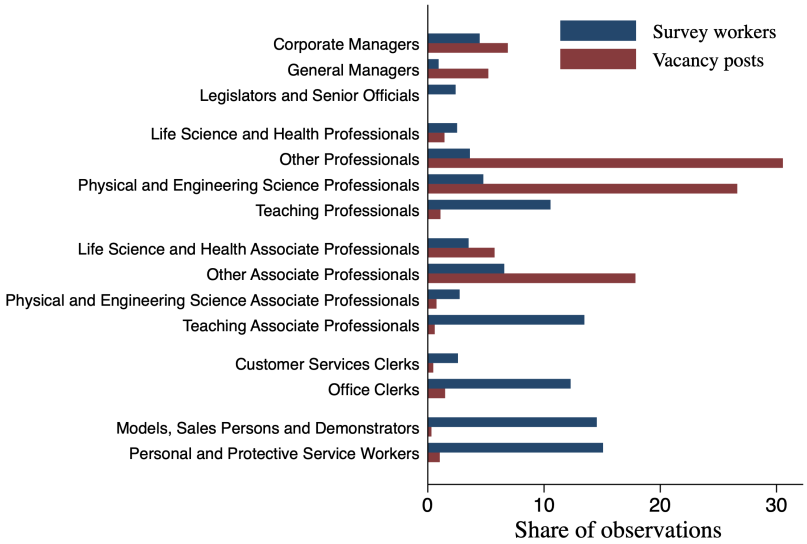
(a) Firm distribution



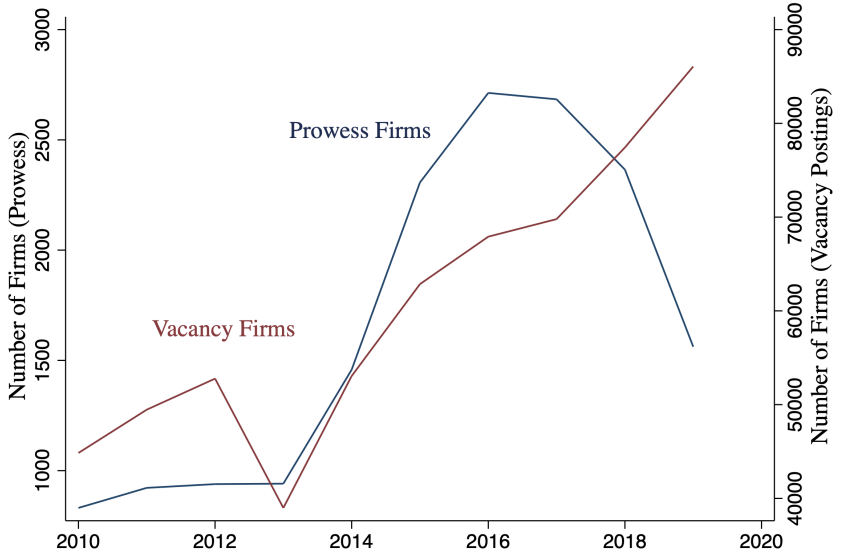
(b) Worker/vacancy distributions



(c) Occupation distribution



(d) Number of firms by year



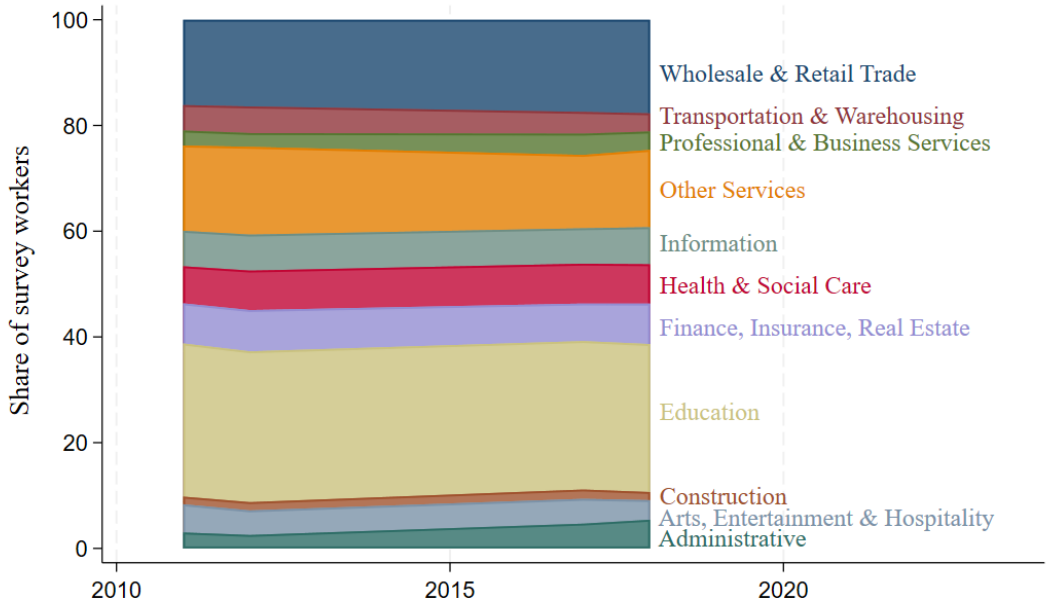
Notes: These figures compare the composition of our vacancy dataset (red) to that of available administrative datasets (blue). Panel (a) shows the distribution of firms across industries relative to Prowess. Panel (b) compares the distribution of vacancies to that of workers in the NSS and PLFS. Panel (c) shows the distribution of white-collar services occupations relative to NSS and PLFS. Panel (d) compares the number of firms in the vacancy data to that in Prowess.

We can also compare the sectoral distribution of vacancies over time to the sectoral distribution of workers in labor force surveys over time. For labor force surveys, Figure C.A.3 plots data from the years when the NSS and PLFS labor force surveys were conducted (2011, 2012, 2017 and 2018). For our vacancies data, we draw on all available years. Panel (a) shows the distribution of survey workers by 2-digit industry and Panel (b) shows the distribution of vacancies by 2-digit industry. Both panels show that industry shares remain largely constant. The discrepancies in the distributions by source also remain roughly constant—e.g. the Information sector is similarly overrepresented in our data relative to the labor force surveys in every year, and the Education sector is similarly underrepresented in our data in every year.

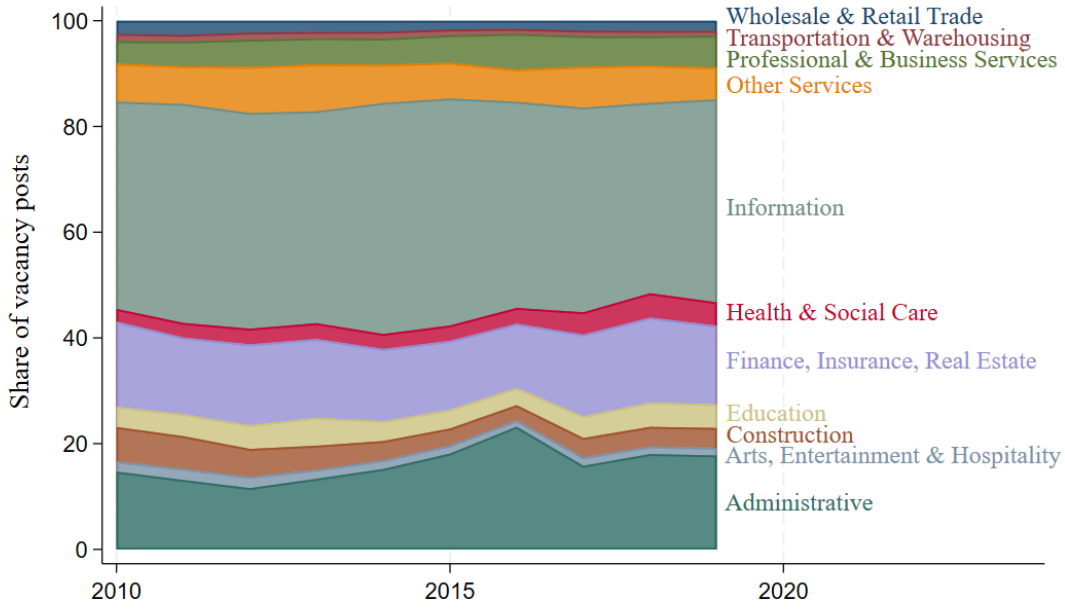
Finally, we examine the firm size distribution in our data by comparing the distribution of vacancies per firm to the distribution of employees per firm in Prowess. Figure C.A.4 shows the results; in both cases, we focus on firms active between 2010 and 2019 and in the services sector. Comparing Panels (a) and (b) illustrates that Prowess is missing the left-tail of the distribution, as expected by its focus on large firms that report financial statements. When analysing employment per firm and vacancies per firm for the subset of firms larger than the median in Prowess (about 3,500 employees), we show in Panels (c) and (d) that our vacancies dataset has a firm size distribution very similar to Prowess.

Figure C.A.3: Industry shares in vacancy data and labor force surveys over time

(a) Sectoral distribution of survey workers, 2-digit industries

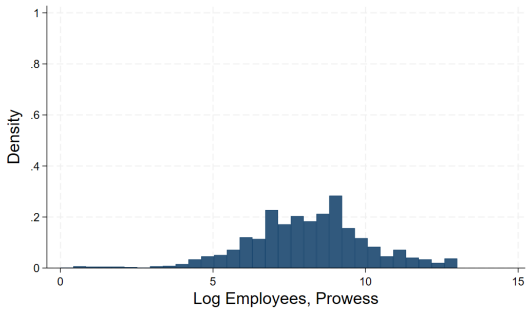


(b) Sectoral distribution of vacancies, 2-digit industries

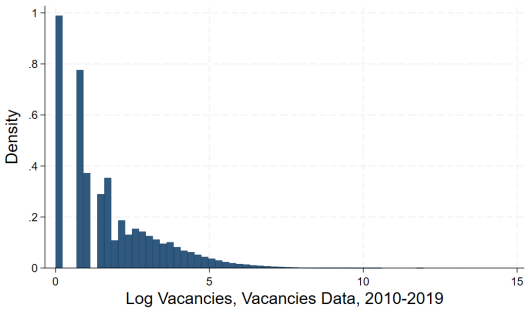


Notes: These figures compare the industry composition of the labor force surveys (NSS and PLFS, top) to that of our vacancy dataset over time (bottom).

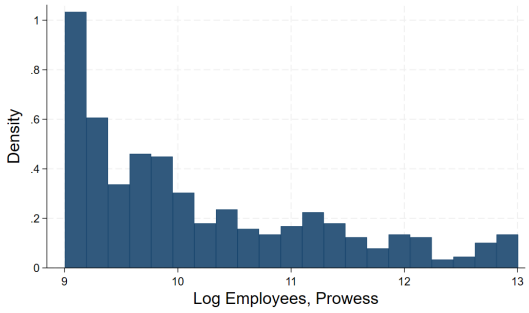
Figure C.A.4: Histogram of employment per firm (Prowess) and vacancies per firm (vacancies data)



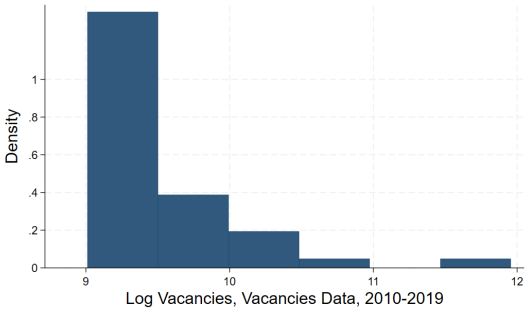
(a) Employment per firm (Prowess)



(b) Vacancies per firm (vacancies data)



(c) Employment per large firm (Prowess)



(d) Vacancies per large firm (vacancies data)

Notes: Panel (a) shows a histogram for employment per firm in Prowess, while Panel (b) shows a histogram for posted vacancies data per firm in our vacancies data. Panel (c) restricts the employment per firm in Prowess to values above the median and Panel (d) restricts vacancies per firm in our vacancies data by the same cut-off, i.e. the firm size median in Prowess. For both histograms, we restrict the sample to the services sector and to firms active in our sample period, 2010-2019.

C.A.3 Details on the restricted sample

This section provides further details on the 'restricted sample', i.e., the sample on which we run our main long-difference regressions. This sample reflects three sets of restrictions. First, as with all analysis in the paper, we focus on non-remote, full-time, permanent jobs located in India in the white-collar services sector. Second, and as with all causal analysis in the paper (i.e., excluding the descriptives in Section 3.3), we focus on 'AI-using' industries only (as opposed to also including 'AI-producing' industries). Finally, and in contrast to the event study in Section 3.6.2, we focus only on establishments that post vacancies in both the baseline (2010-2012) and endline (2017-2019). Table C.A.2 documents and describes the vacancies data at each stage of data restriction.

Table C.A.2: Vacancies data at different stages of data restriction

Stage	Vacancies	Firms	Establishments	Cities	AI share	Mean Salary	Mean Experience
Unrestricted	19,084,725	203,951	973,219	28,586	0.478%	273,883	1.83
Only domestic & non-remote jobs	18,264,142	196,745	912,749	28,478	0.498%	203,196	1.82
Only full time permanent jobs	14,181,508	158,616	652,585	19,831	0.589%	203,526	1.795
Only service sector	14,160,726	158,475	651,457	19,699	0.589%	203,569	1.794
Only AI-using	8,032,318	109,178	482,755	16,335	0.354%	166,528	1.826
Only baseline and endline years	5,152,571	82,642	361,471	11,408	0.422%	177,972	1.839
Only est. posting in both periods	1,533,481	5,605	24,485	442	0.81%	166,519	1.848

Notes: Descriptive statistics and number of observations at different stages of data restriction. The first row describes the vacancies data in their entirety, whereas the final row describes the vacancies data underlying the main sample used in our analysis, which forms first differences.

Table 3.2.1 in the main paper shows summary statistics for the restricted sample (pooling 2010-12 and 2017-19, unless otherwise stated). Table C.A.3 provides further summary statistics for differences between the baseline and endline. Lastly, Figure C.A.5 shows histograms of key variables and of the log of these variables.

Table C.A.3: Difference (2017-19 over 2010-12) descriptives of vacancy data in restricted sample, all jobs and AI jobs only

Difference in	All Jobs	AI Jobs
Mean # Posts	22	.65
90th Percentile # Posts	50	0
Mean AI Share	.17%	-
Mean Annual Salary, Rupees	17,000	23,000
Median Annual Salary, Rupees	16,000	18,000
90th Percentile Annual Salary, Rupees	27,000	77,000
10th Percentile Annual Salary, Rupees	6,000	16,000
Mean Required Work Experience, Years	.15	.15
Median Required Work Experience, Years	.18	.18
Median Postgraduate Share	0	0
90th Percentile Postgraduate Share	22%	71%

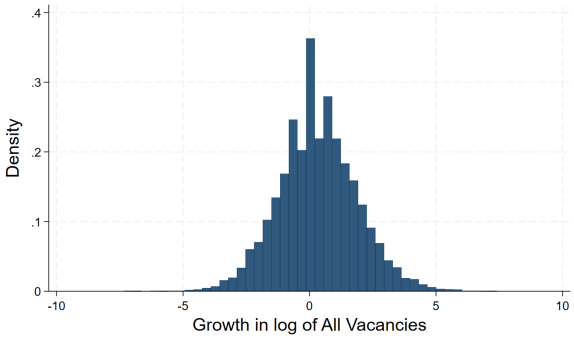
Notes: Descriptive statistics overall and within AI jobs only.

Figure C.A.5: Histograms of key variables

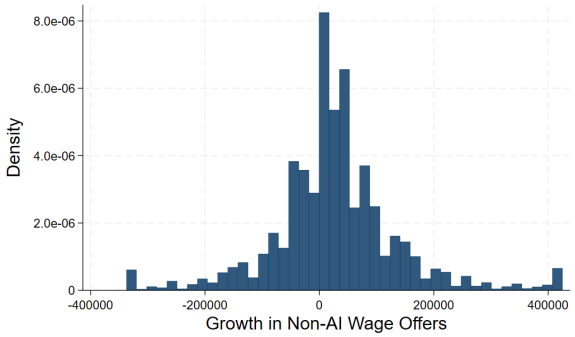
(a) Growth in Total Posts



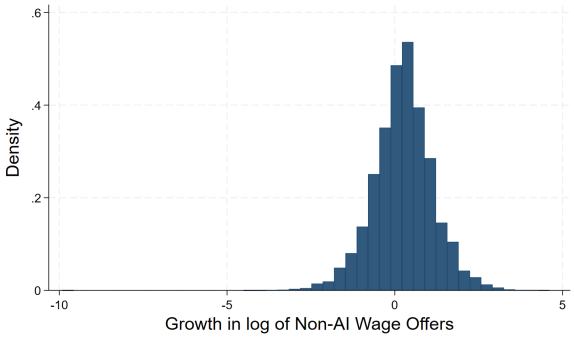
(b) Growth in log of Total Posts



(c) Growth in Non-AI Wage Offers



(d) Growth in log of Non-AI Wage Offers



Notes: This figure shows histograms for growth in vacancies and growth in non-AI wage offers. The left hand side shows growth in levels and the right hand side shows growth in the log of the respective variable. For growth in levels, we display the variables winsorized at the 1 percent level to exclude outliers and improve the legibility of these graphs.

C.A.4 Calculating the AI wage premium

This section provides details on our calculation of the AI wage premium. When including industry-region, industry-time and region-time fixed effects, we find that AI posts on average offer 30 percent higher wages than non-AI posts (see model (1) of Table C.A.4). However, this may be driven by the highest-paying firms also disproportionately hiring AI roles. Therefore, we add firm fixed effects to control for differences between firms in model (2). Even in this case, AI posts pay 19 percent more relative to the average non-AI post. Finally, posts that require AI skills may simply be different types of jobs. Models (3) and (4) therefore include fixed effects for the occupation and role, using respectively the NCO 2004 classification codes and the more granular role label built into the online jobs site. A substantial AI premium of 13-17 percent remains.[4]

[4]The interpretation of the control variables is as follows. An extra year of experience is associated with a more than 35 percent higher salary (at least within the predominantly early-career jobs posted on the site – see Figure 3.3.3), while having a Master's degree is associated with up to 10 percent higher salary. In this sample, having only a high school education is associated with wage offers 3-6 percent below the baseline of having an undergraduate degree, though this figure is likely a dramatic underestimate of the effect, given the major under-representation of lower-skilled professions on the platform. The relationship

Table C.A.4: Wages in AI vs. non-AI roles

	log Annual Salary			
	(1)	(2)	(3)	(4)
AI post	0.318***	0.198***	0.128***	0.174***
	(0.0484)	(0.0358)	(0.0220)	(0.0422)
Experience Required (Years)	0.470***	0.411***	0.386***	0.351***
	(0.00693)	(0.00797)	(0.00787)	(0.00818)
High School	0.00481	-0.0644***	-0.0408**	-0.0395**
	(0.0788)	(0.0192)	(0.0185)	(0.0172)
Master's	0.104***	0.0774***	0.0448***	0.0198**
	(0.0144)	(0.00990)	(0.00781)	(0.00814)
Doctorate	0.131**	0.0741*	0.0132	0.00218
	(0.0588)	(0.0417)	(0.0325)	(0.0339)
Fixed Effects:				
– Industry-Region	✓	✓	✓	✓
– Industry-Year	✓	✓	✓	✓
– Region-Year	✓	✓	✓	✓
– Firm		✓	✓	✓
– Occupation Code			✓	
– Role Label				✓
R^2	0.343	0.535	0.556	0.577
Observations	14012499	13976759	13275348	13976757

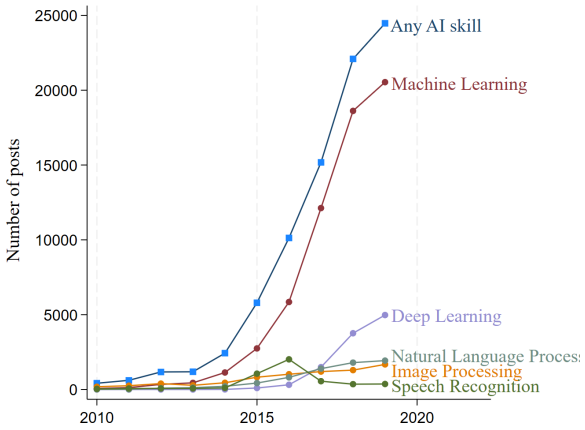
Notes: Standard errors in parentheses. * p <0.10, ** p <0.05, *** p <0.01. All regressions include industry-region, industry-time and region-time fixed effects, and models (2)-(4) also include firm fixed effects. *AI post* is a dummy such that the coefficient is the percentage increase in annual salary associated with posts requiring AI skills, after accounting for the control variables and fixed effects. Similarly, *Experience* is measured in years, so the coefficient reflects the percentage salary increase associated with an additional year of experience. The education variables are dummies, with the baseline category being a Bachelor's degree; for instance, *High School* reflects the percentage salary decrease associated with posts that only require a high school education. The *Occupation Code* fixed effect also accounts for variation across India's 4-digit National Classification of Occupations codes, while the more granular *Role Label* fixed effect accounts for variation across the firm-specified role categories built into the jobs portal. Standard errors clustered at the firm level.

C.A.5 Further descriptives on AI demand and exposure

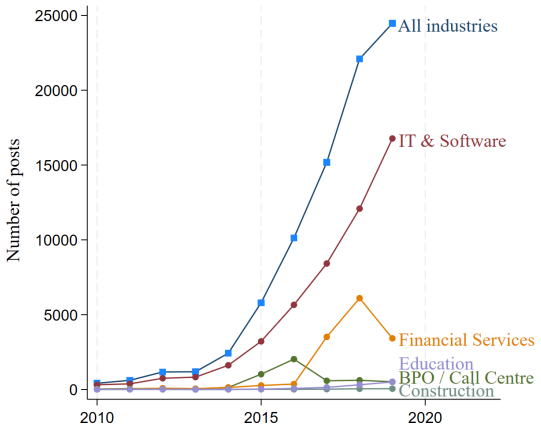
This section collects additional descriptives on AI demand and exposure. As with Section 3.3, these figures all cover both the AI-producing and AI-using sectors. Figure C.A.6 repeats Figure 3.3.1 for the number of posts instead of their share. Figure C.A.7 plots the cumulative share of AI posts against the corresponding share of all posts. The distribution of cities' shares of AI posts and all posts is shown in Figure C.A.8, while the city-wise distribution of AI posts over time is shown in Figure C.A.9. The top 20 roles demanding AI skills in our analysis are listed in Figure C.A.10. Figure C.A.11 decomposes the AI share of job posts by establishment age. Table C.A.5 lists examples of nouns that frequently appear in connection with the top ten verbs that are relatively more common in AI posts, as described Table 3.3.1.[5]

Figure C.A.6: Trends in AI demand, number of posts

(a) Most demanded AI skills



(b) Number of AI posts, by industry

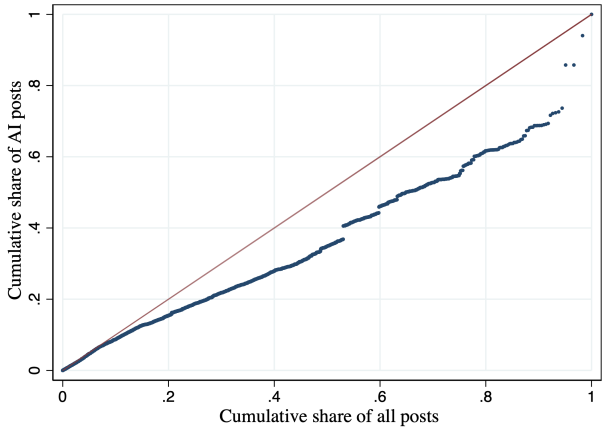


Notes: Panel (a) shows the number of all vacancies that specify particular AI skills, for the top five most demanded skills. Panel (b) shows the number of vacancies that are AI vacancies, both for all industries together and within each of the top five industries by AI share.

between wage offers and having a doctoral degree is expressed predominantly through the firm- and role-effects: conditional on firm and occupation/role, there is no significant relationship to salary, but without such conditioning salaries are 7-13 percent higher. This is consistent with the wage offer premium for workers with doctorates being driven by taking higher-skilled jobs at more advanced firms.

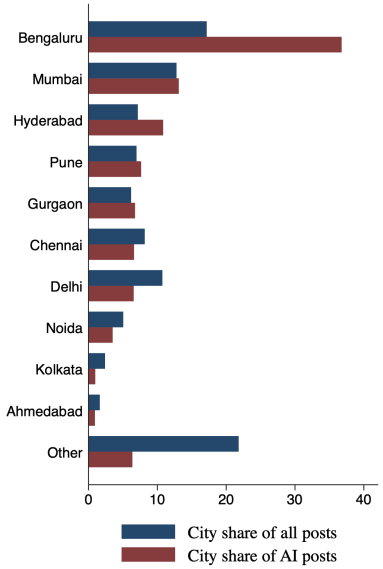
[5]See Table 10 in Webb (2020) for a similar list of the top verbs and example nouns extracted from AI patents (and hence used to construct his exposure measure).

Figure C.A.7: AI posts by firm size



Notes: We plot the cumulative share of AI posts against the corresponding cumulative share of all posts, for the whole period 2010-2019. The red 45° line indicates a one-for-one increase in the share of AI posts relative to all posts. The deviation of our scatter plot from the 45° line shows the extent to which AI vacancies are disproportionately posted by the largest firms.

Figure C.A.8: Shares of posts by city



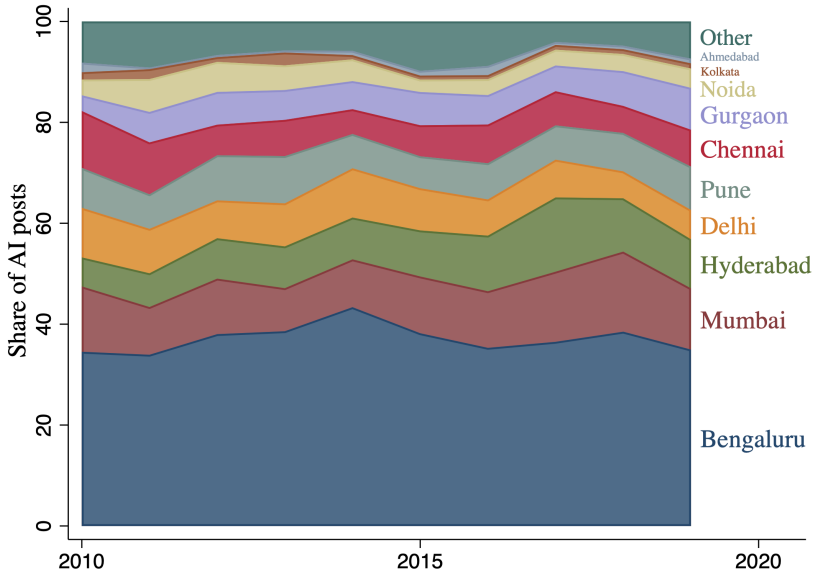
Notes: Bars show the shares of all posts and AI posts across cities, for the entire period 2010 to 2019.

Table C.A.5: Verbs in AI posts and example nouns

Verb	**Example nouns**
Learn	Market Trends, Customer Needs, Framework, Data Analysis, Decision-Making Models
Develop	Solutions, Value Proposition, Automation Tools, Data Pipeline, Insights, Models
Process	Images, Data, Complaints, Requests, Financial Transaction, Files, Natural Language
Program	Software Solution, Process, User Experience, Simulation, System
Predict	Sales, Spending, Growth, Demand, Data Patterns, User Behaviour, Market Changes
Build	Technology, Data Infrastructure, Digital Products, Processes, Roadmap
Work	Data, Digital Transformation, Solution, Problem-Solving
Design	Interface, Strategy, Workflow, Campaign, Algorithm, User Journey, Dashboard
Advance	Business, Career, Company, Credentials, Brand, Operations, Automation Strategies, Technology
Compute	Statistics, Metrics, Averages, Projections, Results, Scores, Costs, ROI

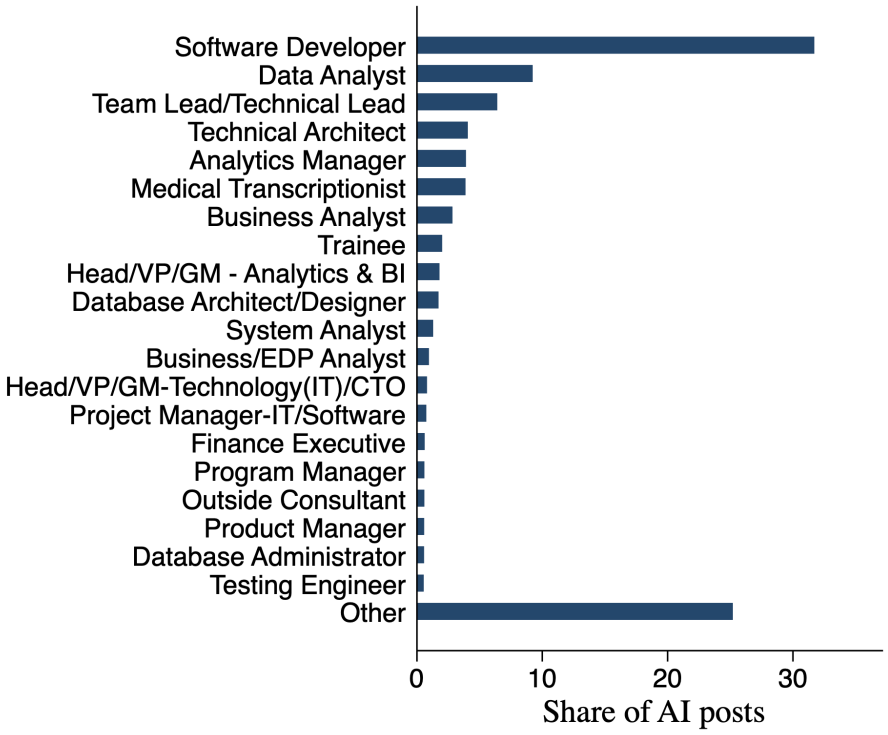
Notes: This table repeats the top ten verbs that are relatively more common in AI posts, as described in Table 3.3.1, and provides examples of the nouns with which they frequently appear.

Figure C.A.9: Cities' shares of AI posts over time



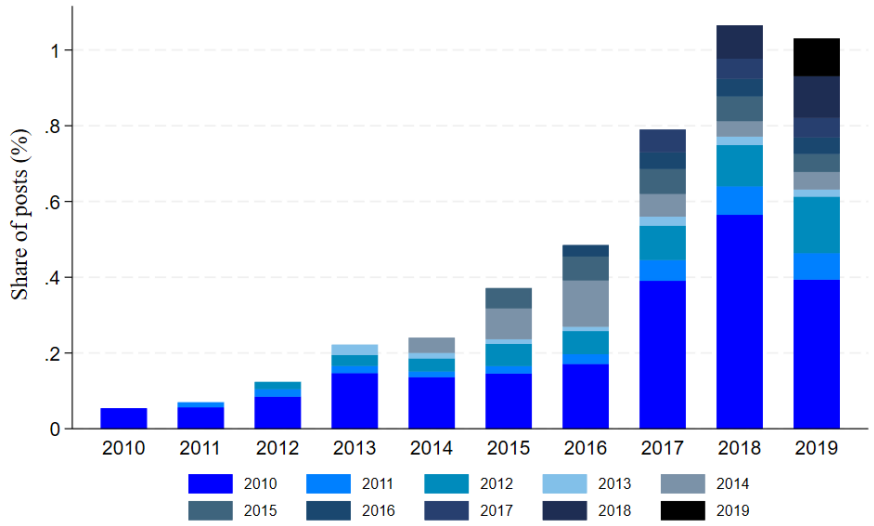
Notes: This graph shows the distribution of AI posts across cities over time. Each year reflects the share of all AI vacancies in that year which were in each city. Shares have been remarkably constant. Bengaluru's share peaked at just over 40 percent in 2014, then Mumbai's share in particular has risen subsequently as AI demand increased in finance (see Figure 3.3.1).

Figure C.A.10: Top 20 roles demanding AI skills, 2010-2019



Notes: We rank the top roles demanding AI skills by their share of AI posts. All other roles hiring AI skills are grouped in the 'Other' category.

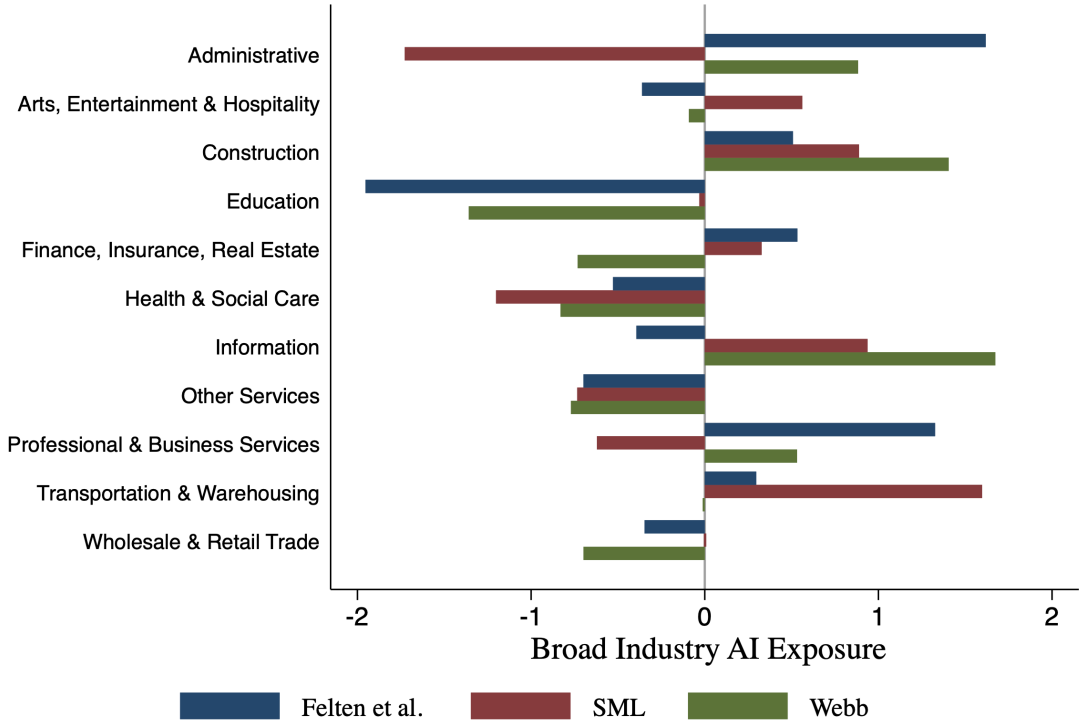
Figure C.A.11: AI share of posts, by establishment age



Notes: We decompose the overall share of posts demanding any AI skill by the year an establishment first posts a vacancy in our data.

Figure C.A.12: AI exposure across industries

(a) AI exposure, by broad industry



Notes: This graph shows standardised AI exposure scores by broad industry categories. The industry exposure scores are a weighted average of occupational exposure, weighted by the vacancy shares of each occupation within the industry. The occupational exposure measures are those of Felten et al. (2018), Mani et al. (2020, building on Brynjolfsson & Mitchell 2017) and Webb (2020) respectively.

Appendix C.B Additional Results and Robustness

This section collects additional results related to our shift-share instrument, our main specification, and various robustness checks, as referenced in the main text.

C.B.1 Shift-share validity and inference

We construct our instrument from baseline (2010-2012) occupation shares at the establishment level and their respective exposure to AI according to Webb (2020):

$$Exposure_{fr,t_0} = \sum_o PostingShare_{fro}^{t_0} \cdot ExposureMeasure_o \tag{C.B.1}$$

This is a Bartik style instrument with occupation shares in the pre-AI baseline that capture an establishment's exposure to a common shock: occupation-level advances in AI. We can test for the exogeneity of the baseline shares following Goldsmith-Pinkham et al. (2020), who propose several validity checks by analogy with GMM and DiD: investigating correlates of shares, examining pre-trends, and comparing different estimators and running over-identification tests. We find that all three provide support for the validity of our instrument.

Test #1: Investigating correlates of shares. We investigate the extent to which the baseline shares correlate with baseline controls, which could themselves affect hiring and wage offer trends. To this end, we regress the instrument on baseline controls (the structure of required education, experience, and wage offers in an establishment). Table C.B.1 shows the results, demonstrating that this does not appear to be an issue for the overall instrument. Some individual occupation shares warrant the inclusion of controls, in particular experience, and we thus confirm robustness to including these controls in our main specification.

Test #2: Examining pre-trends. Most of our results are derived from the long-difference specification discussed above. Therefore, we do not have a pre-period and cannot test for pre-trends. This corresponds to the first empirical example given in Goldsmith-Pinkham et al. (2020), where the shares are fixed in a time period from which we are forming the first difference, such that there is no pre-period. We can, however, ask whether our instrument, which is based on baseline occupation shares, predicts year-on-year hiring or wage growth. We regress annual employment and wage growth from 2014 onwards (so that the first differences do not contain the baseline years, 2010-2012, from whose occupation shares the instrument is constructed) on the instrument. The results are shown in Table C.B.2: we do not find any indication of pre-trends: baseline exposure to AI does not predict differential growth rates. This remains the case when

Table C.B.1: Investigating correlates of shares

	Overall Instrument			
	(1)	(2)	(3)	(4)
Share of Highschool Education	-0.166	0.0203	-0.0615	0.0253
	(0.192)	(0.0929)	(0.146)	(0.0926)
Share of Undergraduate Education	-0.232	0.00812	-0.122	0.0131
	(0.194)	(0.0915)	(0.146)	(0.0912)
Share of Postgraduate Education	-0.221	0.0403	-0.0999	0.0454
	(0.195)	(0.0933)	(0.147)	(0.0928)
Mean Salary	4.86e-09	4.96e-09	4.19e-09	4.97e-09
	(4.59e-09)	(4.85e-09)	(4.37e-09)	(4.87e-09)
Mean Experience	-0.00217	0.00524	0.00334	0.00512
	(0.00823)	(0.00442)	(0.00590)	(0.00445)
Fixed Effects:				
– Region		✓	✓	✓
– Industry		✓		✓
– Firm Decile			✓	✓
Observations	22,201	22,052	22,052	22,052

Notes: Standard errors in parentheses. Standard errors clustered at the firm level. The dependent variable is establishment AI exposure. This is calculated as the standardized average of occupation AI exposure (from Webb 2020), over the occupations for which the establishment posts vacancies in 2010-12, weighted by the number of vacancies posted per occupation, as in Acemoglu et al. (2022). The independent variables are baseline controls.

including the set of fixed effects included in our main regressions.

Table C.B.2: Examining pre-trends for the instrument

	Growth in Non-AI Vacancies				Growth in Non-AI Median Wage			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Instrument	0.749	-0.492	-0.101	-0.495	-116.1	-439.5	-88.87	-438.3
	(2.273)	(0.703)	(1.885)	(0.703)	(611.1)	(464.0)	(616.1)	(463.1)
Fixed Effects:								
– Region		✓	✓	✓		✓	✓	✓
– Industry		✓		✓		✓		✓
– Firm Decile			✓	✓			✓	✓
Observations	296,730	296,730	296,730	296,730	296,730	296,730	296,730	296,730

Notes: Standard errors in parentheses. Standard errors clustered at the firm level. The independent variable is establishment AI exposure. This is calculated as the standardized average of occupation AI exposure (from Webb 2020), over the occupations for which the establishment posts vacancies in 2010-12, weighted by the number of vacancies posted per occupation, as in Acemoglu et al. (2022). The dependent variables are year-on year growth for 2013-2019.

Test #3: Alternative estimators and over-identification tests. We next compare a range of estimators (OLS, a range of IV estimators, a machine learning estimator and a Fuller-like estimator) and run over-identification tests. Following Goldsmith-Pinkham et al. (2020), we compare Bartik to OLS, over-identified TSLS, using each share as a separate instrument, the Modified Bias-corrected TSLS (MBTSLS) estimator, the Limited Information Maximum Likelihood (LIML) estimator, and the HFUL estimator. Similarity in results between HFUL and LIML on the one hand, and MBTSLS and over-identified TSLS on the other hand supports the validity of our instrument. Bartik estimates are similar to LIML estimates when including establishment controls. Results from HFUL and MBTSLS are also similar, further supporting our instrument. The comparison of alternative estimators suggests validity of our instrument as we find estimates to be quite similar.

We then run over-identification tests for the HFUL, LIML, and over-identified TSLS estimators, where the null hypothesis is the validity of the over-identifying restrictions. These tests do not reject the null hypothesis when including controls. For misspecification tests, we test whether Bartik is sensitive to the inclusion of controls. Similarity in estimates would support our instrument, and indeed we find support for our instrument's validity.

Adjusted standard errors. In addition to validity, a further issue with shift-share instruments concerns standard errors that are correlated. Table C.B.3 presents results when instead computing standard errors according to the correction developed by Adão et al. (2019), and finds that our results are robust.

Table C.B.3: Second stage: Impact of AI adoption on establishment non-AI vacancies. Adão, Kolesár, and Morales (2019) standard errors.

	Growth in Non-AI Vacancies			Growth in Total Vacancies		
	(1)	(2)	(3)	(4)	(5)	(6)
Adoption of AI	-7.975**	-12.90*	-8.064***	-7.737**	-12.47*	-7.840***
	(3.654)	(7.549)	(3.265)	(3.659)	(7.564)	(3.269)
Fixed Effects:						
– Region	✓	✓	✓	✓	✓	✓
– Industry	✓		✓	✓		✓
– Firm Decile		✓	✓		✓	✓
First Stage F-Stat	26.06	26.31	27.17	26.06	26.31	27.17
Observations	22,251	22,251	22,251	22,251	22,251	22,251

Notes: Standard errors calculated as in Adão, Kolesár, and Morales (2019) in parentheses. * p <0.10, ** p <0.05, *** p <0.01. Standard errors clustered at the firm level. The independent variable is a dummy, which equals 1 for establishments that did not post AI vacancies in the baseline period, but posted AI vacancies in the endline period. The dependent variables are the change in the log of the respective establishment-level outcomes, specifically the number of non-AI vacancies and the total number of vacancies. AI adoption is instrumented by establishment AI exposure. Each coefficient therefore represents the percentage point impact upon the outcome variable of a one percent rise in the predicted probability of adopting AI. AI exposure is calculated as the standardized average of occupation AI exposure (from Webb 2020), over the occupations for which the establishment posts vacancies in 2010-12, weighted by the number of vacancies posted per occupation, as in Acemoglu et al. (2022).

C.B.2 Additional results from main specification

This section presents additional results from our main specification. As this concerns the causal identification of the effects of AI demand on labor demand and wage offers, we are excluding AI-producing sectors from this analysis, and keep only white-collar services jobs in AI-using industries. In Table C.B.4, we show the first stage. Tables C.B.5 and C.B.6 repeat the main regressions on hiring and wage offers when dropping the ten largest firms. Tables C.B.7 and C.B.8 repeat the main regressions when dropping the three establishments accounting for most of the 2016-17 growth in AI vacancies. Similarly, Tables C.B.9 and C.B.10 drop the 5 establishments accounting for almost two thirds of AI job posts as per Table 3.2.1. Tables C.B.11 and C.B.12 drop the top decile of adopters instead. Tables C.B.13 and C.B.14 repeat the main regressions when excluding establishments in the finance sector. We show that this sector, which is a major user of AI in India, is not solely driving our findings.

Figure C.B.1 extends the wage distribution graph to all postings. Table C.B.15 studies the impact of AI on 2-digit occupations.[6] Table C.B.16 shows task growth results for abstract and routine tasks following Autor & Dorn (2013) and confirms the findings of Table 3.5.3. Similarly,

[6]The estimated effect differs significantly across occupation groups, with the hypothesis of equal coefficients rejected at the 1 percent level (p value=0.00). Differences are also generally statistically significant for pairwise comparisons at the 1 percent level, but insignificant for the comparisons of 'General Managers' and 'Other Professionals', and of 'Health Professionals' and 'Teaching Professionals'.

Table C.B.17 repeats Table 3.5.3 while controlling for changes in occupational shares. Table C.B.18 repeats Table 3.5.4 for abstract tasks following Autor & Dorn (2013). Table C.B.19 shows wage growth results by 1-digit occupations.[7] Finally, Table C.B.20 shows wage growth results when controlling for job profiles.

Table C.B.4: First stage: Impact of AI exposure on establishment AI adoption

	Adoption of AI		
	(1)	(2)	(3)
Establishment AI Exposure	0.0103***	0.00965***	0.0106***
	(0.00156)	(0.00149)	(0.00157)
Fixed Effects:			
– Region	✓	✓	✓
– Firm Decile		✓	✓
– Industry	✓		✓
R^2	.0547	.0434	.062
Observations	22,251	22,251	22,251

Notes: Standard errors in parentheses. * p <0.10, ** p <0.05, *** p <0.01. Standard errors clustered at the firm level. The dependent variable is a dummy, which equals 1 for establishments that did not post AI vacancies in the baseline period, but posted AI vacancies in the endline period. The independent variable is establishment AI exposure, calculated as the standardized average of occupation AI exposure (from Webb 2020), over the occupations for which the establishment posts vacancies in 2010-12, weighted by the number of vacancies posted per occupation, as in Acemoglu et al. (2022). Each coefficient therefore represents the impact on the probability of adopting AI of a one-standard deviation rise in AI exposure according to a linear probability model.

[7]The estimated effect does not differ significantly across occupation groups (p value=0.0949).

Table C.B.5: Second stage: Impact of AI adoption on establishment non-AI vacancies, excluding ten largest firms

	Growth in Non-AI Vacancies			Growth in Total Vacancies		
	(1)	(2)	(3)	(4)	(5)	(6)
Adoption of AI	-8.297***	-13.38***	-8.339***	-8.041***	-12.91***	-8.098***
	(2.366)	(3.187)	(2.304)	(2.259)	(3.044)	(2.199)
Fixed Effects:						
– Region	✓	✓	✓	✓	✓	✓
– Industry	✓		✓	✓		✓
– Firm Decile		✓	✓		✓	✓
First Stage F-Stat	43.01	39.73	44.35	43.38	40.02	44.57
Observations	21,595	21,595	21,595	21,602	21,602	21,602

Notes: Standard errors in parentheses. * p <0.10, ** p <0.05, *** p <0.01. Standard errors clustered at the firm level. The independent variable is a dummy, which equals 1 for establishments that did not post AI vacancies in the baseline period, but posted AI vacancies in the endline period. The dependent variables are the change in the log of the respective establishment-level outcomes, specifically the number of non-AI vacancies and the total number of vacancies. AI adoption is instrumented by establishment AI exposure. Each coefficient therefore represents the percentage point impact upon the outcome variable of a one percent rise in the predicted probability of adopting AI. AI exposure is calculated as the standardized average of occupation AI exposure (from Webb 2020), over the occupations for which the establishment posts vacancies in 2010-12, weighted by the number of vacancies posted per occupation, as in Acemoglu et al. (2022). The ten largest firms in terms of baseline (2010-2012) hiring are excluded.

Table C.B.6: Second stage: Impact of AI adoption on establishment non-AI wages, excluding ten largest firms

	Growth in Non-AI Median Wage			Growth in Overall Median Wage		
	(1)	(2)	(3)	(4)	(5)	(6)
Adoption of AI	-5.549***	-6.287***	-5.399***	-5.307***	-6.019***	-5.157***
	(1.483)	(1.661)	(1.431)	(1.422)	(1.593)	(1.373)
Fixed Effects:						
– Region	✓	✓	✓	✓	✓	✓
– Industry	✓		✓	✓		✓
– Firm Decile		✓	✓		✓	✓
First Stage F-Stat	44.44	41.01	45.76	44.78	41.27	45.94
Observations	21,417	21,417	21,417	21,424	21,424	21,424

Notes: Standard errors in parentheses. * p <0.10, ** p <0.05, *** p <0.01. Standard errors clustered at the firm level. The independent variable is a dummy, which equals 1 for establishments that did not post AI vacancies in the baseline period, but posted AI vacancies in the endline period. The dependent variables are the change in the log of the respective establishment-level outcomes, specifically wages of non-AI vacancies and wages of all vacancies. AI adoption is instrumented by establishment AI exposure. Each coefficient therefore represents the percentage point impact upon the outcome variable of a one percent rise in the predicted probability of adopting AI. AI exposure is calculated as the standardized average of occupation AI exposure (from Webb 2020), over the occupations for which the establishment posts vacancies in 2010-12, weighted by the number of vacancies posted per occupation, as in Acemoglu et al. (2022). The ten largest firms in terms of baseline (2010-2012) hiring are excluded.

Table C.B.7: Second stage: Impact of AI adoption on establishment non-AI vacancies, excluding the three establishments accounting for most of the 2016-17 growth in AI vacancies

	Growth in Non-AI Vacancies			Growth in Total Vacancies		
	(1)	(2)	(3)	(4)	(5)	(6)
Adoption of AI	-8.205***	-13.23***	-8.295***	-7.958***	-12.78***	-8.062***
	(2.411)	(3.176)	(2.340)	(2.301)	(3.036)	(2.233)
Fixed Effects:						
– Region	✓	✓	✓	✓	✓	✓
– Industry	✓		✓	✓		✓
– Firm Decile		✓	✓		✓	✓
First Stage F-Stat	42.7	40.61	44.43	43.05	40.87	44.62
Observations	22,241	22,241	22,241	22,248	22,248	22,248

Notes: Standard errors in parentheses. * p <0.10, ** p <0.05, *** p <0.01. Standard errors clustered at the firm level. The independent variable is a dummy, which equals 1 for establishments that did not post AI vacancies in the baseline period, but posted AI vacancies in the endline period. The dependent variables are the change in the log of the respective establishment-level outcomes, specifically the number of non-AI vacancies and the total number of vacancies. AI adoption is instrumented by establishment AI exposure. Each coefficient therefore represents the percentage point impact upon the outcome variable of a one percent rise in the predicted probability of adopting AI. AI exposure is calculated as the standardized average of occupation AI exposure (from Webb 2020), over the occupations for which the establishment posts vacancies in 2010-12, weighted by the number of vacancies posted per occupation, as in Acemoglu et al. (2022). The three establishments accounting for most of the 2016-17 growth in AI vacancies are excluded.

Table C.B.8: Second stage: Impact of AI adoption on establishment non-AI wages, excluding the three establishments accounting for most of the 2016-17 growth in AI vacancies

	Growth in Non-AI Median Wage			Growth in Overall Median Wage		
	(1)	(2)	(3)	(4)	(5)	(6)
Adoption of AI	-5.854***	-6.513***	-5.664***	-5.599***	-6.239***	-5.411***
	(1.522)	(1.671)	(1.458)	(1.459)	(1.604)	(1.398)
Fixed Effects:						
– Region	✓	✓	✓	✓	✓	✓
– Industry	✓		✓	✓		✓
– Firm Decile		✓	✓		✓	✓
First Stage F-Stat	44.13	41.93	45.85	44.45	42.15	46
Observations	22,061	22,061	22,061	22,068	22,068	22,068

Notes: Standard errors in parentheses. * p <0.10, ** p <0.05, *** p <0.01. Standard errors clustered at the firm level. The independent variable is a dummy, which equals 1 for establishments that did not post AI vacancies in the baseline period, but posted AI vacancies in the endline period. The dependent variables are the change in the log of the respective establishment-level outcomes, specifically wages of non-AI vacancies and wages of all vacancies. AI adoption is instrumented by establishment AI exposure. Each coefficient therefore represents the percentage point impact upon the outcome variable of a one percent rise in the predicted probability of adopting AI. AI exposure is calculated as the standardized average of occupation AI exposure (from Webb 2020), over the occupations for which the establishment posts vacancies in 2010-12, weighted by the number of vacancies posted per occupation, as in Acemoglu et al. (2022). The three establishments accounting for most of the 2016-17 growth in AI vacancies are excluded.

Table C.B.9: Second stage: Impact of AI adoption on establishment non-AI vacancies, excluding the 5 establishments contributing two thirds of AI job posts

	Growth in Non-AI Vacancies			Growth in Total Vacancies		
	(1)	(2)	(3)	(4)	(5)	(6)
Adoption of AI	-8.233***	-13.28***	-8.327***	-7.985***	-12.83***	-8.093***
	(2.415)	(3.187)	(2.344)	(2.305)	(3.046)	(2.237)
Fixed Effects:						
– Region	✓	✓	✓	✓	✓	✓
– Industry	✓		✓	✓		✓
– Firm Decile		✓	✓		✓	✓
First Stage F-Stat	42.6	40.5	44.32	42.95	40.75	44.51
Observations	22,239	22,239	22,239	22,246	22,246	22,246

Notes: Standard errors in parentheses. * p <0.10, ** p <0.05, *** p <0.01. Standard errors clustered at the firm level. The independent variable is a dummy, which equals 1 for establishments that did not post AI vacancies in the baseline period, but posted AI vacancies in the endline period. The dependent variables are the change in the log of the respective establishment-level outcomes, specifically the number of non-AI vacancies and the total number of vacancies. AI adoption is instrumented by establishment AI exposure. Each coefficient therefore represents the percentage point impact upon the outcome variable of a one percent rise in the predicted probability of adopting AI. AI exposure is calculated as the standardized average of occupation AI exposure (from Webb 2020), over the occupations for which the establishment posts vacancies in 2010-12, weighted by the number of vacancies posted per occupation, as in Acemoglu et al. (2022). The 5 establishments contributing two thirds of AI job posts as per Appendix Table 3.2.1 have been excluded.

Table C.B.10: Second stage: Impact of AI adoption on establishment non-AI wages, excluding the 5 establishments contributing two thirds of AI job posts

	Growth in Non-AI Median Wage			Growth in Overall Median Wage		
	(1)	(2)	(3)	(4)	(5)	(6)
Adoption of AI	-5.864***	-6.531***	-5.674***	-5.608***	-6.256***	-5.420***
	(1.525)	(1.676)	(1.461)	(1.462)	(1.609)	(1.401)
Fixed Effects:						
– Region	✓	✓	✓	✓	✓	✓
– Industry	✓		✓	✓		✓
– Firm Decile		✓	✓		✓	✓
First Stage F-Stat	44.02	41.81	45.74	44.35	42.03	45.9
Observations	22,059	22,059	22,059	22,066	22,066	22,066

Notes: Standard errors in parentheses. * p <0.10, ** p <0.05, *** p <0.01. Standard errors clustered at the firm level. The independent variable is a dummy, which equals 1 for establishments that did not post AI vacancies in the baseline period, but posted AI vacancies in the endline period. The dependent variables are the change in the log of the respective establishment-level outcomes, specifically wages of non-AI vacancies and wages of all vacancies. AI adoption is instrumented by establishment AI exposure. Each coefficient therefore represents the percentage point impact upon the outcome variable of a one percent rise in the predicted probability of adopting AI. AI exposure is calculated as the standardized average of occupation AI exposure (from Webb 2020), over the occupations for which the establishment posts vacancies in 2010-12, weighted by the number of vacancies posted per occupation, as in Acemoglu et al. (2022). The 5 establishments contributing two thirds of AI job posts as per Appendix Table 3.2.1 have been excluded.

Table C.B.11: Second stage: Impact of AI adoption on establishment non-AI vacancies, excluding top decile of firms by adoption

	Growth in Non-AI Vacancies			Growth in Total Vacancies		
	(1)	(2)	(3)	(4)	(5)	(6)
Adoption of AI	-12.08***	-20.05***	-12.53***	-11.59***	-19.15***	-12.05***
	(3.438)	(4.965)	(3.357)	(3.270)	(4.722)	(3.196)
Fixed Effects:						
– Region	✓	✓	✓	✓	✓	✓
– Industry	✓		✓	✓		✓
– Firm Decile		✓	✓		✓	✓
First Stage F-Stat	33.98	29.18	35.19	33.13	28.71	34.19
Observations	21,183	21,183	21,183	21,188	21,188	21,188

Notes: Standard errors in parentheses. * p <0.10, ** p <0.05, *** p <0.01. Standard errors clustered at the firm level. The independent variable is a dummy, which equals 1 for establishments that did not post AI vacancies in the baseline period, but posted AI vacancies in the endline period. The dependent variables are the change in the log of the respective establishment-level outcomes, specifically the number of non-AI vacancies and the total number of vacancies. AI adoption is instrumented by establishment AI exposure. Each coefficient therefore represents the percentage point impact upon the outcome variable of a one percent rise in the predicted probability of adopting AI. AI exposure is calculated as the standardized average of occupation AI exposure (from Webb 2020), over the occupations for which the establishment posts vacancies in 2010-12, weighted by the number of vacancies posted per occupation, as in Acemoglu et al. (2022). The top decile of firms by adoption, measured by the growth in AI vacancies posted between the endline and baseline, is excluded.

Table C.B.12: Second stage: Impact of AI adoption on establishment non-AI wages, excluding top decile of firms by adoption

	Growth in Non-AI Median Wage			Growth in Overall Median Wage		
	(1)	(2)	(3)	(4)	(5)	(6)
Adoption of AI	-8.012***	-9.366***	-7.738***	-7.603***	-8.886***	-7.341***
	(2.203)	(2.602)	(2.114)	(2.110)	(2.486)	(2.028)
Fixed Effects:						
– Region	✓	✓	✓	✓	✓	✓
– Industry	✓		✓	✓		✓
– Firm Decile		✓	✓		✓	✓
First Stage F-Stat	34.21	29.32	35.4	33.35	28.85	34.4
Observations	21,006	21,006	21,006	21,011	21,011	21,011

Notes: Standard errors in parentheses. * p <0.10, ** p <0.05, *** p <0.01. Standard errors clustered at the firm level. The independent variable is a dummy, which equals 1 for establishments that did not post AI vacancies in the baseline period, but posted AI vacancies in the endline period. The dependent variables are the change in the log of the respective establishment-level outcomes, specifically wages of non-AI vacancies and wages of all vacancies. AI adoption is instrumented by establishment AI exposure. Each coefficient therefore represents the percentage point impact upon the outcome variable of a one percent rise in the predicted probability of adopting AI. AI exposure is calculated as the standardized average of occupation AI exposure (from Webb 2020), over the occupations for which the establishment posts vacancies in 2010-12, weighted by the number of vacancies posted per occupation, as in Acemoglu et al. (2022). The top decile of firms by adoption, measured by the growth in AI vacancies posted between the endline and baseline, is excluded.

Table C.B.13: Second stage: Impact of AI adoption on establishment non-AI vacancies, excluding finance sector

	Growth in Non-AI Vacancies			Growth in Total Vacancies		
	(1)	(2)	(3)	(4)	(5)	(6)
Adoption of AI	-14.42***	-17.58***	-14.47***	-14.65***	-17.85***	-14.70***
	(4.008)	(4.233)	(3.950)	(4.112)	(4.349)	(4.056)
Fixed Effects:						
– Region	✓	✓	✓	✓	✓	✓
– Industry	✓		✓	✓		✓
– Firm Decile		✓	✓		✓	✓
First Stage F-Stat	28.19	32.57	28.82	27.27	31.49	27.81
Observations	16,558	16,558	16,558	16,561	16,561	16,561

Notes: Standard errors in parentheses. * p <0.10, ** p <0.05, *** p <0.01. Standard errors clustered at the firm level. The independent variable is a dummy, which equals 1 for establishments that did not post AI vacancies in the baseline period, but posted AI vacancies in the endline period. The dependent variables are the change in the log of the respective establishment-level outcomes, specifically the number of non-AI vacancies and the total number of vacancies. AI adoption is instrumented by establishment AI exposure. Each coefficient therefore represents the percentage point impact upon the outcome variable of a one percent rise in the predicted probability of adopting AI. AI exposure is calculated as the standardized average of occupation AI exposure (from Webb 2020), over the occupations for which the establishment posts vacancies in 2010-12, weighted by the number of vacancies posted per occupation, as in Acemoglu et al. (2022). Establishments in the finance sector have been excluded.

Table C.B.14: Second stage: Impact of AI adoption on establishment non-AI wages, excluding finance sector

	Growth in Non-AI Median Wage			Growth in Overall Median Wage		
	(1)	(2)	(3)	(4)	(5)	(6)
Adoption of AI	-8.346***	-8.257***	-8.217***	-8.477***	-8.378***	-8.344***
	(2.436)	(2.285)	(2.377)	(2.496)	(2.339)	(2.437)
Fixed Effects:						
– Region	✓	✓	✓	✓	✓	✓
– Industry	✓		✓	✓		✓
– Firm Decile		✓	✓		✓	✓
First Stage F-Stat	28.32	32.47	28.9	27.39	31.39	27.9
Observations	16,421	16,421	16,421	16,424	16,424	16,424

Notes: Standard errors in parentheses. * p <0.10, ** p <0.05, *** p <0.01. Standard errors clustered at the firm level. The independent variable is a dummy, which equals 1 for establishments that did not post AI vacancies in the baseline period, but posted AI vacancies in the endline period. The dependent variables are the change in the log of the respective establishment-level outcomes, specifically wages of non-AI vacancies and wages of all vacancies. AI adoption is instrumented by establishment AI exposure. Each coefficient therefore represents the percentage point impact upon the outcome variable of a one percent rise in the predicted probability of adopting AI. AI exposure is calculated as the standardized average of occupation AI exposure (from Webb 2020), over the occupations for which the establishment posts vacancies in 2010-12, weighted by the number of vacancies posted per occupation, as in Acemoglu et al. (2022). Establishments in the finance sector have been excluded.

Table C.B.15: Second stage: Impact of AI adoption on establishment non-AI vacancy shares, by granular occupation group

	Change in Non-AI Vacancy Shares					
	Engineering Professionals	Health Professionals	Teaching Professionals	Other Professionals	Corporate Managers	General Managers
Adoption of AI	-2.689***	0.130	0.212***	-1.290***	-9.964***	-0.626**
	(0.494)	(0.120)	(0.0748)	(0.409)	(1.589)	(0.299)
Fixed Effects:						
– Region	✓	✓	✓	✓	✓	✓
– Industry	✓	✓	✓	✓	✓	✓
– Firm Decile	✓	✓	✓	✓	✓	✓
First Stage F-Stat	45.43	45.43	45.43	45.43	45.43	45.43
Observations	22,244	22,244	22,244	22,244	22,244	22,244

Notes: Standard errors in parentheses. * p <0.10, ** p <0.05, *** p <0.01. Standard errors clustered at the firm level. The independent variable is a dummy, which equals 1 for establishments that did not post AI vacancies in the baseline period, but posted AI vacancies in the endline period. The dependent variables are the change in occupation shares. Occupation groups are 2-digit occupations within Professionals and Managers from the NCO04. AI adoption is instrumented by establishment AI exposure. Each coefficient therefore represents the impact upon the outcome variable of a one percent rise in the predicted probability of adopting AI. AI exposure is calculated as the standardized average of occupation AI exposure (from Webb 2020), over the occupations for which the establishment posts vacancies in 2010-12, weighted by the number of vacancies posted per occupation, as in Acemoglu et al. (2022).

Table C.B.16: Second stage: Impact of AI adoption on establishment abstract and routine tasks

	Growth in Abstract Tasks			Growth in Routine Tasks		
	(1)	(2)	(3)	(4)	(5)	(6)
Adoption of AI	-5.753***	-7.208***	-5.772***	-0.219	-0.193	-0.217
	(1.094)	(1.348)	(1.073)	(0.214)	(0.193)	(0.213)
Fixed Effects:						
– Region	✓	✓	✓	✓	✓	✓
– Industry	✓		✓	✓		✓
– Firm Decile		✓	✓		✓	✓
First Stage F-Stat	44.06	41.83	45.62	44.06	41.83	45.62
Observations	22,251	22,251	22,251	22,251	22,251	22,251

Notes: Standard errors in parentheses. * p <0.10, ** p <0.05, *** p <0.01. Standard errors clustered at the firm level. The independent variable is a dummy, which equals 1 for establishments that did not post AI vacancies in the baseline period, but posted AI vacancies in the endline period. The dependent variables are the change in the log of the respective establishment-level outcomes, specifically routine and abstract tasks, where we have shifted the mean of the distributions of scores to be able to take logarithms. We use the occupational task scores for abstract, routine, and manual tasks from Autor & Dorn (2013) (based on data from the Dictionary of Occupational Titles 1977) and map occ1990dd occupations to NCO04 occupations. Scores are standardized as in Acemoglu & Autor (2011). AI adoption is instrumented by establishment AI exposure. Each coefficient therefore represents the percentage point impact upon the outcome variable of a one percent rise in the predicted probability of adopting AI. AI exposure is calculated as the standardized average of occupation AI exposure (from Webb 2020), over the occupations for which the establishment posts vacancies in 2010-12, weighted by the number of vacancies posted per occupation, as in Acemoglu et al. (2022).

Table C.B.17: Second stage: Impact of AI adoption on establishment routine and non-routine tasks, holding occupational composition fixed

	Growth in Non-Routine Tasks			Growth in Routine Tasks		
	(1)	(2)	(3)	(4)	(5)	(6)
Adoption of AI	-3.632***	-3.450***	-3.481***	0.124	0.0911	0.117
	(0.918)	(0.908)	(0.874)	(0.856)	(0.851)	(0.817)
Fixed Effects:						
– Region	✓	✓	✓	✓	✓	✓
– Industry	✓		✓	✓		✓
– Firm Decile		✓	✓		✓	✓
First Stage F-Stat	51.82	58.16	55.11	51.82	58.16	55.11
Observations	22,251	22,251	22,251	22,251	22,251	22,251

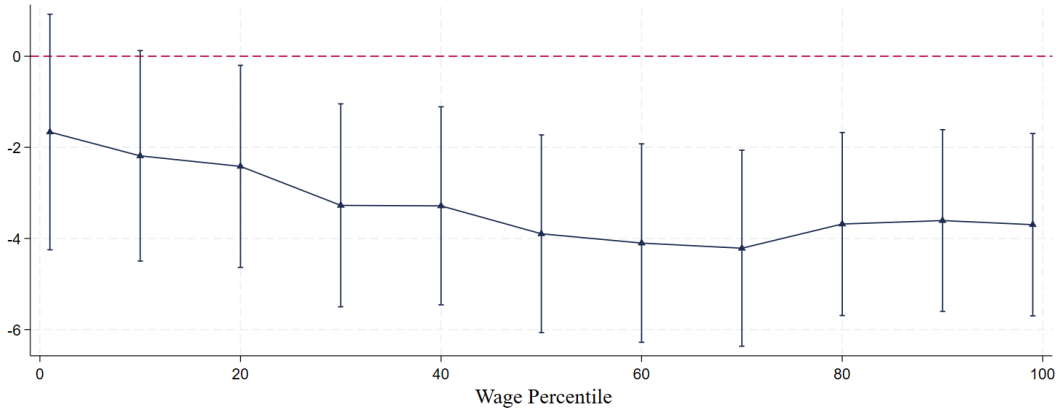
Notes: Standard errors in parentheses. * p <0.10, ** p <0.05, *** p <0.01. Standard errors clustered at the firm level. The independent variable is a dummy, which equals 1 for establishments that did not post AI vacancies in the baseline period, but posted AI vacancies in the endline period. We control for the change in the log of total posts and for the change in 1-digit NCO04 occupations, leaving out . The dependent variables are the change in the log of the respective establishment-level outcomes, where we have shifted the mean of the distributions of routine and non-routine scores to be able to take logarithms. We average standardised routine and non-routine O*NET task contents by occupation, and form establishments' routine and non-routine task demand by weighting occupations by their standardised routine and non-routine scores. Each coefficient therefore represents the percentage point impact upon the outcome variable of a one percent rise in the predicted probability of adopting AI. AI exposure is calculated as the standardized average of occupation AI exposure (from Webb 2020), over the occupations for which the establishment posts vacancies in 2010-12, weighted by the number of vacancies posted per occupation, as in Acemoglu et al. (2022).

Table C.B.18: Second stage: Impact of AI adoption on establishment abstract tasks, by occupation group

	Growth in Abstract Tasks				
	Personal, Sales & Security	Clerks	Associate Professionals	Professionals	Managers
Adoption of AI	-0.0477**	-0.00733	0.296***	-1.651**	-29.15***
	(0.0225)	(0.0121)	(0.114)	(0.780)	(4.863)
Fixed Effects:					
– Region	✓	✓	✓	✓	✓
– Industry	✓	✓	✓	✓	✓
– Firm Decile	✓	✓	✓	✓	✓
First Stage F-Stat	45.62	45.62	45.62	45.62	45.62
Observations	22,251	22,251	22,251	22,251	22,251

Notes: Standard errors in parentheses. * p <0.10, ** p <0.05, *** p <0.01. Standard errors clustered at the firm level. The independent variable is a dummy, which equals 1 for establishments that did not post AI vacancies in the baseline period, but posted AI vacancies in the endline period. The dependent variables are the change in the log of the respective establishment-level outcomes, specifically abstract tasks, where we have shifted the mean of the distributions of scores to be able to take logarithms. We use the occupational task scores for abstract, routine, and manual tasks from Autor & Dorn (2013) (based on data from the Dictionary of Occupational Titles 1977) and map occ1990dd occupations to NCO04 occupations. Scores are standardized as in Acemoglu & Autor (2011). AI adoption is instrumented by establishment AI exposure. Each coefficient therefore represents the percentage point impact upon the outcome variable of a one percent rise in the predicted probability of adopting AI. AI exposure is calculated as the standardized average of occupation AI exposure (from Webb 2020), over the occupations for which the establishment posts vacancies in 2010-12, weighted by the number of vacancies posted per occupation, as in Acemoglu et al. (2022).

Figure C.B.1: Impact of AI adoption on the wage offer distribution in all posts



Notes: This coefficient plot shows the impact of a one percent rise in the predicted probability of adopting AI on wage growth over time across the distribution of establishment wage offers. Each coefficient is from a regression of type (6) in Appendix Table C.B.20. As in Appendix Table C.B.20, AI adoption is instrumented with AI exposure. In other words, each coefficient represents the impact of a one percent rise in the predicted probability of adopting AI on wage growth over time for a given percentile of the wage offer distribution. We report the 1st and 99th percentile of the wage offer distribution and deciles in between the two extremes, alongside 95 percent confidence intervals. Standard errors are clustered at the firm level, and we include region, firm size decile and industry fixed effects. Since AI posts make up only a small share of all roles in most establishments, the pattern is very similar across the distributions for all posts and for non-AI posts only.

Table C.B.19: Second stage: Impact of AI adoption on establishment non-AI wages, by occupation group

	Growth in Non-AI Median Wage				
	(1)	(2)	(3)	(4)	(5)
Adoption of AI	0.716	1.816*	-0.840	-1.073	-2.037*
	(2.060)	(1.067)	(0.729)	(0.942)	(1.194)
Fixed Effects:					
– Region	✓	✓	✓	✓	✓
– Industry	✓	✓	✓	✓	✓
– Firm Decile	✓	✓	✓	✓	✓
First Stage F-Stat	6.443	11.67	61.28	23.57	25.89
Observations	981	2,059	13,128	9,296	8,003

Notes: Standard errors in parentheses. * p <0.10, ** p <0.05, *** p <0.01. Standard errors clustered at the firm level. The independent variable is a dummy, which equals 1 for establishments that did not post AI vacancies in the baseline period, but posted AI vacancies in the endline period. The dependent variables are the change in the log of the respective establishment-level outcomes, specifically wages of non-AI vacancies and wages of all vacancies. AI adoption is instrumented by establishment AI exposure. Occupation groups are 1-digit occupation groups from the NCO04. Each coefficient therefore represents the percentage point impact upon the outcome variable of a one percent rise in the predicted probability of adopting AI. AI exposure is calculated as the standardized average of occupation AI exposure (from Webb 2020), over the occupations for which the establishment posts vacancies in 2010-12, weighted by the number of vacancies posted per occupation, as in Acemoglu et al. (2022).

Table C.B.20: Second stage: Impact of AI adoption dummy on establishment non-AI wages, controlling for job profiles

	Growth in Non-AI Median Wage			Growth in Overall Median Wage		
	(1)	(2)	(3)	(4)	(5)	(6)
Adoption of AI	-4.147***	-4.347***	-4.049***	-3.997***	-4.196***	-3.898***
	(1.191)	(1.278)	(1.152)	(1.144)	(1.229)	(1.107)
Growth in Experience	0.409***	0.409***	0.405***	0.409***	0.409***	0.405***
	(0.0132)	(0.0138)	(0.0133)	(0.0132)	(0.0136)	(0.0132)
Growth in High School share	-0.0883	-0.0896	-0.102	-0.0919	-0.0928	-0.105
	(0.0724)	(0.0740)	(0.0716)	(0.0705)	(0.0720)	(0.0697)
Growth in Master's share	0.222***	0.220***	0.222***	0.221***	0.219***	0.222***
	(0.0320)	(0.0318)	(0.0318)	(0.0320)	(0.0318)	(0.0317)
Growth in Doctorate share	2.090**	2.283**	2.045**	2.029**	2.219**	1.985**
	(0.988)	(1.085)	(0.968)	(0.957)	(1.052)	(0.937)
Fixed Effects:						
– Region	✓	✓	✓	✓	✓	✓
– Industry	✓		✓	✓		✓
– Firm Decile		✓	✓		✓	✓
First Stage F-Stat	45.86	43.58	47.26	46.13	43.73	47.38
Observations	22,064	22,064	22,064	22,071	22,071	22,071

Notes: Standard errors in parentheses. * p <0.10, ** p <0.05, *** p <0.01. Standard errors clustered at the firm level. The independent variable is a dummy, which equals 1 for establishments that did not post AI vacancies in the baseline period, but posted AI vacancies in the endline period. The dependent variables are the change in the log of the respective establishment-level outcomes, specifically wages of non-AI vacancies and wages of all vacancies. AI adoption is instrumented by establishment AI exposure. Each coefficient therefore represents the percentage point impact upon the outcome variable of a one percent rise in the predicted probability of adopting AI. AI exposure is calculated as the standardized average of occupation AI exposure (from Webb 2020), over the occupations for which the establishment posts vacancies in 2010-12, weighted by the number of vacancies posted per occupation, as in Acemoglu et al. (2022).

C.B.3 Alternative specifications

This section provides the key results repeated for a series of alternative specifications, as discussed in Section 3.6, along with additional results for the event-study approach.

Table C.B.21: Second stage: Impact of AI adoption on establishment non-AI vacancies, controlling for baseline share of software engineers

	Growth in Non-AI Vacancies			Growth in Total Vacancies		
	(1)	(2)	(3)	(4)	(5)	(6)
Adoption of AI	-9.656***	-16.10***	-9.582***	-9.353***	-15.51***	-9.293***
	(2.944)	(4.206)	(2.865)	(2.801)	(3.997)	(2.725)
Covariates:						
Share of Software Engineers	✓	✓	✓	✓	✓	✓
Fixed Effects:						
– Region	✓	✓	✓	✓	✓	✓
– Industry	✓		✓	✓		✓
– Firm Decile		✓	✓		✓	✓
First Stage F-Stat	32.3	28.77	33.05	33.01	29.4	33.69
Observations	22,244	22,244	22,244	22,251	22,251	22,251

Notes: Standard errors in parentheses. * p <0.10, ** p <0.05, *** p <0.01. Standard errors clustered at the firm level. The independent variable is a dummy, which equals 1 for establishments that did not post AI vacancies in the baseline period, but posted AI vacancies in the endline period. The dependent variables are the change in the log of the respective establishment-level outcomes, specifically the number of non-AI vacancies and the total number of vacancies. AI adoption is instrumented by establishment AI exposure. As an additional control, the baseline share of vacancies for software engineers in each establishment is included. Software engineers are captured by the NCO04 4-digit occupations 2131 (Computer Systems Designers and Analysts), 2132 (Computer Programmers), and 2139 (Computer Professionals, n.e.c.). Each coefficient therefore represents the percentage point impact upon the outcome variable of a one percent rise in the predicted probability of adopting AI. AI exposure is calculated as the standardized average of occupation AI exposure (from Webb 2020), over the occupations for which the establishment posts vacancies in 2010-12, weighted by the number of vacancies posted per occupation, as in Acemoglu et al. (2022).

Table C.B.22: Second stage: Impact of AI adoption on establishment non-AI wages, controlling for baseline share of software engineers

	Growth in Non-AI Median Wage			Growth in Overall Median Wage		
	(1)	(2)	(3)	(4)	(5)	(6)
Adoption of AI	-6.620***	-7.642***	-6.482***	-6.326***	-7.307***	-6.186***
	(1.833)	(2.121)	(1.776)	(1.749)	(2.024)	(1.695)
Covariates:						
Share of Software Engineers	✓	✓	✓	✓	✓	✓
Fixed Effects:						
– Region	✓	✓	✓	✓	✓	✓
– Industry	✓		✓	✓		✓
– Firm Decile		✓	✓		✓	✓
First Stage F-Stat	33.58	29.9	34.32	34.27	30.52	34.94
Observations	22,064	22,064	22,064	22,071	22,071	22,071

Notes: Standard errors in parentheses. * p <0.10, ** p <0.05, *** p <0.01. Standard errors clustered at the firm level. The independent variable is a dummy, which equals 1 for establishments that did not post AI vacancies in the baseline period, but posted AI vacancies in the endline period. The dependent variables are the change in the log of the respective establishment-level outcomes, specifically wages of non-AI vacancies and wages of all vacancies. AI adoption is instrumented by establishment AI exposure. As an additional control, the baseline share of vacancies for software engineers in each establishment is included. Software engineers are captured by the NCO04 4-digit occupations 2131 (Computer Systems Designers and Analysts), 2132 (Computer Programmers), and 2139 (Computer Professionals, n.e.c.). Each coefficient therefore represents the percentage point impact upon the outcome variable of a one percent rise in the predicted probability of adopting AI. AI exposure is calculated as the standardized average of occupation AI exposure (from Webb 2020), over the occupations for which the establishment posts vacancies in 2010-12, weighted by the number of vacancies posted per occupation, as in Acemoglu et al. (2022).

Table C.B.23: Second stage: Impact of AI adoption on establishment non-AI vacancies, controlling for baseline share of sales & admin vacancies

	Growth in Non-AI Vacancies			Growth in Total Vacancies		
	(1)	(2)	(3)	(4)	(5)	(6)
Adoption of AI	-17.96**	-40.93**	-18.08***	-16.82***	-36.83**	-16.96***
	(6.976)	(20.67)	(6.916)	(6.335)	(17.49)	(6.270)
Fixed Effects:						
– Region	✓	✓	✓	✓	✓	✓
– Industry	✓		✓	✓		✓
– Firm Decile		✓	✓		✓	✓
First Stage F-Stat	10.49	4.644	10.69	11.17	5.282	11.35
Observations	22,244	22,244	22,244	22,251	22,251	22,251

Notes: Standard errors in parentheses. * p <0.10, ** p <0.05, *** p <0.01. Standard errors clustered at the firm level. The independent variable is a dummy, which equals 1 for establishments that did not post AI vacancies in the baseline period, but posted AI vacancies in the endline period. The dependent variables are the change in the log of the respective establishment-level outcomes, specifically the number of non-AI vacancies and the total number of vacancies. As an additional control, the baseline share of vacancies in each establishment belonging to the broad occupations of either sales or administration is included. We use the occ1990dd occupation classification (by Autor & Dorn 2013) in defining retail sales and clerical jobs as sales and administrative. Each coefficient therefore represents the percentage point impact upon the outcome variable of a one percent rise in the predicted probability of adopting AI. AI exposure is calculated as the standardized average of occupation AI exposure (from Webb 2020), over the occupations for which the establishment posts vacancies in 2010-12, weighted by the number of vacancies posted per occupation, as in Acemoglu et al. (2022).

Table C.B.24: Second stage: Impact of AI adoption on establishment non-AI wages, controlling for baseline share of sales & admin vacancies

	Growth in Non-AI Median Wage			Growth in Overall Median Wage		
	(1)	(2)	(3)	(4)	(5)	(6)
Adoption of AI	-6.393**	-10.85*	-6.318**	-6.393**	-10.85*	-6.318**
	(3.005)	(6.168)	(2.933)	(3.005)	(6.168)	(2.933)
Fixed Effects:						
– Region	✓	✓	✓	✓	✓	✓
– Industry	✓		✓	✓		✓
– Firm Decile		✓	✓		✓	✓
First Stage F-Stat	10.48	4.705	10.76	10.48	4.705	10.76
Observations	22,064	22,064	22,064	22,064	22,064	22,064

Notes: Standard errors in parentheses. * p <0.10, ** p <0.05, *** p <0.01. Standard errors clustered at the firm level. The independent variable is a dummy, which equals 1 for establishments that did not post AI vacancies in the baseline period, but posted AI vacancies in the endline period. The dependent variables are the change in the log of the respective establishment-level outcomes, specifically wages of non-AI vacancies and wages of all vacancies. As an additional control, the baseline share of vacancies in each establishment belonging to the broad occupations of either sales or administration is included. We use the occ1990dd occupation classification (by Autor & Dorn 2013) in defining retail sales and clerical jobs as sales and administrative. Each coefficient therefore represents the percentage point impact upon the outcome variable of a one percent rise in the predicted probability of adopting AI. AI exposure is calculated as the standardized average of occupation AI exposure (from Webb 2020), over the occupations for which the establishment posts vacancies in 2010-12, weighted by the number of vacancies posted per occupation, as in Acemoglu et al. (2022).

Table C.B.25: Second stage: Impact of AI adoption on establishment non-AI vacancies, 2013-15 to 2017-19

	Growth in Non-AI Vacancies			Growth in Total Vacancies		
	(1)	(2)	(3)	(4)	(5)	(6)
Adoption of AI	-6.735**	-10.40***	-7.029**	-6.682**	-10.38***	-6.989**
	(2.878)	(3.404)	(2.782)	(2.916)	(3.445)	(2.816)
Fixed Effects:						
– Region	✓	✓	✓	✓	✓	✓
– Industry	✓		✓	✓		✓
– Firm Decile		✓	✓		✓	✓
First Stage F-Stat	41.13	45.05	44.32	39.45	43.08	42.62
Observations	38,458	38,458	38,458	38,490	38,490	38,490

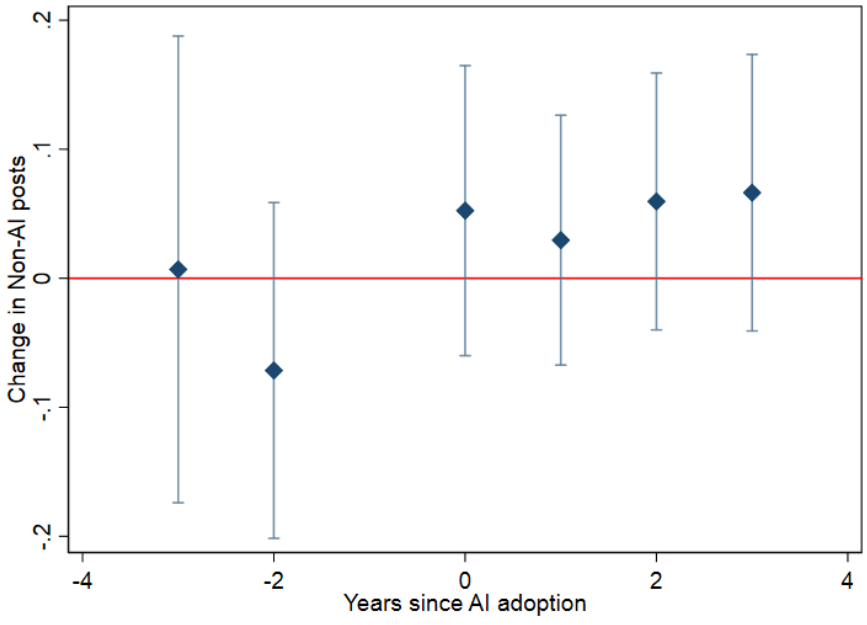
Notes: Standard errors in parentheses. * p <0.10, ** p <0.05, *** p <0.01. Standard errors clustered at the firm level. The independent variable is a dummy, which equals 1 for establishments that did not post AI vacancies in the 2013-15 period, but posted AI vacancies in the endline period. The dependent variables are the change in the log of the respective establishment-level outcomes, specifically the number of non-AI vacancies and the total number of vacancies. AI adoption is instrumented by establishment AI exposure. Each coefficient therefore represents the percentage point impact upon the outcome variable of a one percent rise in the predicted probability of adopting AI. AI exposure is calculated as the standardized average of occupation AI exposure (from Webb 2020), over the occupations for which the establishment posts vacancies in 2013-15, weighted by the number of vacancies posted per occupation, as in Acemoglu et al. (2022).

Table C.B.26: Second stage: Impact of AI adoption on establishment non-AI wages, 2013-15 to 2017-19

	Growth in Non-AI Median Wage			Growth in Overall Median Wage		
	(1)	(2)	(3)	(4)	(5)	(6)
Adoption of AI	-7.114***	-6.920***	-6.829***	-7.185***	-6.966***	-6.883***
	(2.002)	(2.015)	(1.911)	(2.038)	(2.045)	(1.940)
Fixed Effects:						
– Region	✓	✓	✓	✓	✓	✓
– Industry	✓		✓	✓		✓
– Firm Decile		✓	✓		✓	✓
First Stage F-Stat	41.17	45.17	44.28	39.49	43.19	42.57
Observations	38,249	38,249	38,249	38,281	38,281	38,281

Notes: Standard errors in parentheses. * p <0.10, ** p <0.05, *** p <0.01. Standard errors clustered at the firm level. The independent variable is a dummy, which equals 1 for establishments that did not post AI vacancies in the 2013-15 period, but posted AI vacancies in the endline period. The dependent variables are the change in the log of the respective establishment-level outcomes, specifically wages of non-AI vacancies and wages of all vacancies. AI adoption is instrumented by establishment AI exposure. Each coefficient therefore represents the percentage point impact upon the outcome variable of a one percent rise in the predicted probability of adopting AI. AI exposure is calculated as the standardized average of occupation AI exposure (from Webb 2020), over the occupations for which the establishment posts vacancies in 2013-15, weighted by the number of vacancies posted per occupation, as in Acemoglu et al. (2022).

Figure C.B.2: Non-AI vacancies following AI adoption at the district level



Notes: Two way fixed effects on a balanced panel. The outcome variable is the IHS-transformed number of non-AI vacancies by region and district, respectively. At the region level, propensity scores from a probit regression on lagged total hiring, lagged median salary growth, firm age, and year dummies for 2013, 2015, 2016, and 2018. At the district level, propensity scores from a probit regression on lagged median salary, firm age, median salary growth, and year dummies for 2023 and 2013. We use three leads and lags, leaving out the first lead (t-1) as the base period, and cluster standard errors on region and district, respectively. AI adoption leads to reduced non-AI vacancies on the region level, but has no effects on the district level.

Table C.B.27: Vacancies results for 'incumbents' and 'startups'

	Growth in Non-AI Vacancies			Growth in Total Vacancies		
	(1)	(2)	(3)	(4)	(5)	(6)
Adoption of AI	-5.420***	-6.204***	-5.162***	-5.392***	-6.202***	-5.147***
Incumbents	(2.065)	(2.171)	(1.938)	(2.087)	(2.200)	(1.960)
Fixed Effects:						
– Region	✓	✓	✓	✓	✓	✓
– Industry	✓		✓	✓		✓
– Firm Decile		✓	✓		✓	✓
First Stage F-Stat	34.59	35.31	38.14	33.72	34.22	37.11
Observations	17,341	17,341	17,341	17,348	17,348	17,348
Adoption of AI	-11.20	-21.30*	-12.30	-11.17	-21.26*	-12.29
Start-ups	(8.499)	(11.96)	(8.430)	(8.689)	(12.15)	(8.605)
Fixed Effects:						
– Region	✓	✓	✓	✓	✓	✓
– Industry	✓		✓	✓		✓
– Firm Decile		✓	✓		✓	✓
First Stage F-Stat	9.133	9.889	9.725	8.54	9.245	9.138
Observations	21,060	21,060	21,060	21,085	21,085	21,085

Notes: Standard errors in parentheses. * p <0.10, ** p <0.05, *** p <0.01. Standard errors clustered at the firm level. The independent variable is a dummy, which equals 1 for establishments that did not post AI vacancies in the baseline period, but posted AI vacancies in the endline period. The dependent variables are the change in the log of the respective establishment-level outcomes, specifically the number of non-AI vacancies and the total number of vacancies. In order to distinguish between start-ups and incumbents, we look at the shorter long difference between 2013-15 and 2017-19. A start-up is an establishment that did not post in the baseline, 2010-12, and only started posting in 2013-15. An incumbent posted vacancies already in the baseline, 2010-12. AI adoption is instrumented by establishment AI exposure. Each coefficient therefore represents the percentage point impact upon the outcome variable of a one percent rise in the predicted probability of adopting AI. AI exposure is calculated as the standardized average of occupation AI exposure (from Webb 2020), over the occupations for which the establishment posts vacancies in 2013-15, weighted by the number of vacancies posted per occupation, as in Acemoglu et al. (2022).

Table C.B.28: Wage results for 'incumbents' and 'startups'

	Growth in Non-AI Median Wage			Growth in Overall Median Wage		
	(1)	(2)	(3)	(4)	(5)	(6)
Adoption of AI	-3.853***	-3.596***	-3.650***	-3.913***	-3.655***	-3.703***
Incumbent	(1.250)	(1.189)	(1.163)	(1.267)	(1.207)	(1.178)
Fixed Effects:						
– Region	✓	✓	✓	✓	✓	✓
– Industry	✓		✓	✓		✓
– Firm Decile		✓	✓		✓	✓
First Stage F-Stat	34.71	35.28	38.23	33.84	34.19	37.19
Observations	17,259	17,259	17,259	17,266	17,266	17,266
Adoption of AI	-16.27**	-17.74**	-15.81**	-16.48**	-17.74**	-15.97**
Start-ups	(7.103)	(7.778)	(6.760)	(7.347)	(7.949)	(6.966)
Fixed Effects:						
– Region	✓	✓	✓	✓	✓	✓
– Industry	✓		✓	✓		✓
– Firm Decile		✓	✓		✓	✓
First Stage F-Stat	9.135	10.08	9.757	8.538	9.432	9.161
Observations	20,934	20,934	20,934	20,959	20,959	20,959

Notes: Standard errors in parentheses. * p <0.10, ** p <0.05, *** p <0.01. Standard errors clustered at the firm level. The independent variable is a dummy, which equals 1 for establishments that did not post AI vacancies in the baseline period, but posted AI vacancies in the endline period. The dependent variables are the change in the log of the respective establishment-level outcomes, specifically wages of non-AI vacancies and wages of all vacancies. In order to distinguish between start-ups and incumbents, we look at the shorter long difference between 2013-15 and 2017-19. A start-up is an establishment that did not post in the baseline, 2010-12, and only started posting in 2013-15. An incumbent posted vacancies already in the baseline, 2010-12. AI adoption is instrumented by establishment AI exposure. Each coefficient therefore represents the percentage point impact upon the outcome variable of a one percent rise in the predicted probability of adopting AI. AI exposure is calculated as the standardized average of occupation AI exposure (from Webb 2020), over the occupations for which the establishment posts vacancies in 2013-15, weighted by the number of vacancies posted per occupation, as in Acemoglu et al. (2022).

Table C.B.29: Second stage: Impact of AI adoption on establishment non-AI vacancies, intensive margin of AI demand

	Growth in Non-AI Vacancies			Growth in Total Vacancies		
	(1)	(2)	(3)	(4)	(5)	(6)
Growth of AI Vacancies	-4.611***	-7.704***	-4.632***	-4.555***	-7.564***	-4.584***
	(1.557)	(2.187)	(1.517)	(1.513)	(2.115)	(1.475)
Fixed Effects:						
– Region	✓	✓	✓	✓	✓	✓
– Industry	✓		✓	✓		✓
– Firm Decile		✓	✓		✓	✓
First Stage F-Stat	22.76	23.26	23.74	23.9	24.34	24.9
Observations	22,244	22,244	22,244	22,251	22,251	22,251

Notes: Standard errors in parentheses. * p <0.10, ** p <0.05, *** p <0.01. Standard errors clustered at the firm level. The independent variable is the growth in establishment AI vacancies between 2010-12 and 2017-19, approximated by the change in the the logarithm plus 1. The dependent variables are the change in the log of the respective establishment-level outcomes, specifically the number of non-AI vacancies and the total number of vacancies. AI adoption is instrumented by establishment AI exposure. AI vacancies growth is instrumented by establishment AI exposure. AI exposure is calculated as the standardized average of occupation AI exposure (from Webb 2020), over the occupations for which the establishment posts vacancies in 2010-12, weighted by the number of vacancies posted per occupation, as in Acemoglu et al. (2022).

Table C.B.30: Second stage: Impact of AI dummy on establishment non-AI wages, intensive margin of AI demand

	Growth in Non-AI Median Wage			Growth in Overall Median Wage		
	(1)	(2)	(3)	(4)	(5)	(6)
Growth of AI Vacancies	-3.302***	-3.803***	-3.174***	-3.218***	-3.703***	-3.090***
	(0.997)	(1.119)	(0.946)	(0.962)	(1.079)	(0.912)
Fixed Effects:						
– Region	✓	✓	✓	✓	✓	✓
– Industry	✓		✓	✓		✓
– Firm Decile		✓	✓		✓	✓
First Stage F-Stat	23.22	23.71	24.18	24.37	24.8	25.36
Observations	22,064	22,064	22,064	22,071	22,071	22,071

Notes: Standard errors in parentheses. * p <0.10, ** p <0.05, *** p <0.01. Standard errors clustered at the firm level. The independent variable is the growth in establishment AI vacancies between 2010-12 and 2017-19, approximated by the change in the the logarithm plus 1. The dependent variables are the change in the log of the respective establishment-level outcomes, specifically wages of non-AI vacancies and wages of all vacancies. AI adoption is instrumented by establishment AI exposure. AI vacancies growth is instrumented by establishment AI exposure. AI exposure is calculated as the standardized average of occupation AI exposure (from Webb 2020), over the occupations for which the establishment posts vacancies in 2010-12, weighted by the number of vacancies posted per occupation, as in Acemoglu et al. (2022).

Table C.B.31: Second stage: Impact of AI adoption on establishment non-AI vacancies, IHS of AI vacancies

	Growth in Non-AI Vacancies			Growth in Total Vacancies		
	(1)	(2)	(3)	(4)	(5)	(6)
Growth in AI Vacancies	-3.777***	-6.285***	-3.794***	-3.727***	-6.165***	-3.751***
	(1.252)	(1.745)	(1.219)	(1.215)	(1.686)	(1.184)
Fixed Effects:						
– Region	✓	✓	✓	✓	✓	✓
– Industry	✓		✓	✓		✓
– Firm Decile		✓	✓		✓	✓
First Stage F-Stat	24.79	25.12	25.87	26.06	26.31	27.17
Observations	22,244	22,244	22,244	22,251	22,251	22,251

Notes: Standard errors in parentheses. * p <0.10, ** p <0.05, *** p <0.01. Standard errors clustered at the firm level. The independent variable is the growth in establishment AI vacancies between 2010-12 and 2017-19, approximated by the change in the log. The dependent variables are the change in the inverse hyperbolic sine of the respective establishment-level outcomes. AI adoption is instrumented by establishment AI exposure. Each coefficient therefore represents the percentage point impact on the outcome variable of a one IHS unit increase in establishment AI demand. AI exposure is calculated as the standardized average of occupation AI exposure (from Webb 2020), over the occupations for which the establishment posts vacancies in 2010-12, weighted by the number of vacancies posted per occupation, as in Acemoglu et al. (2022).

Table C.B.32: Second stage: Impact of AI adoption on establishment non-AI vacancies, IHS of AI vacancies

	Growth in Non-AI Median Wage			Growth in Overall Median Wage		
	(1)	(2)	(3)	(4)	(5)	(6)
Growth in AI Vacancies	-2.703***	-3.101***	-2.599***	-2.632***	-3.017***	-2.527***
	(0.799)	(0.895)	(0.758)	(0.770)	(0.862)	(0.730)
Fixed Effects:						
– Region	✓	✓	✓	✓	✓	✓
– Industry	✓		✓	✓		✓
– Firm Decile		✓	✓		✓	✓
First Stage F-Stat	25.32	25.64	26.39	26.61	26.84	27.71
Observations	22,064	22,064	22,064	22,071	22,071	22,071

Notes: Standard errors in parentheses. * p <0.10, ** p <0.05, *** p <0.01. Standard errors clustered at the firm level. The independent variable is the growth in establishment AI vacancies between 2010-12 and 2017-19, approximated by the change in the log. The dependent variables are the change in the inverse hyperbolic sine of the respective establishment-level outcomes. AI adoption is instrumented by establishment AI exposure. Each coefficient therefore represents the percentage point impact on the outcome variable of a one IHS unit increase in establishment AI demand. AI exposure is calculated as the standardized average of occupation AI exposure (from Webb 2020), over the occupations for which the establishment posts vacancies in 2010-12, weighted by the number of vacancies posted per occupation, as in Acemoglu et al. (2022).

Table C.B.33: Second stage: Impact of AI adoption on establishment non-AI vacancies and wages, including firm fixed effects

	Non-AI Vacancies	Total Vacancies	Growth in Non-AI Median Wage		Growth in Overall Median Wage	
	(1)	(2)	(3)	(4)	(5)	(6)
Adoption of AI	-18.60**	-18.48**	-10.16**	-8.583**	-10.03**	-8.483**
	(8.988)	(8.946)	(4.961)	(4.282)	(4.914)	(4.243)
Growth in Experience				0.346***		0.346***
				(0.0229)		(0.0228)
Growth in High School share				-0.200***		-0.200***
				(0.0671)		(0.0668)
Growth in Master's share				0.284***		0.284***
				(0.0635)		(0.0638)
Growth in Doctorate share				2.100		2.067
				(1.790)		(1.769)
Fixed Effects:						
– Region	✓	✓	✓	✓	✓	✓
– Firm	✓	✓	✓	✓	✓	✓
First Stage F-Stat	5.538	5.559	6.221	6.045	6.244	6.074
Observations	19,377	19,383	19,217	19,217	19,223	19,223

Notes: Standard errors in parentheses. * p <0.10, ** p <0.05, *** p <0.01. Standard errors clustered at the firm level. The independent variable is a dummy, which equals 1 for establishments that did not post AI vacancies in the baseline period, but posted AI vacancies in the endline period. The dependent variables are the change in the log of the respective establishment-level outcomes, specifically the number of non-AI vacancies, the total number of vacancies, wages of non-AI vacancies and wages of all vacancies. AI adoption is instrumented by establishment AI exposure. Each coefficient therefore represents the percentage point impact upon the outcome variable of a one percent rise in the predicted probability of adopting AI. AI exposure is calculated as the standardized average of occupation AI exposure (from Webb 2020), over the occupations for which the establishment posts vacancies in 2010-12, weighted by the number of vacancies posted per occupation, as in Acemoglu et al. (2022).

Table C.B.34: Second stage: Impact of AI adoption on establishment non-AI vacancies, excluding last year of data

	Growth in Non-AI Vacancies			Growth in Total Vacancies		
	(1)	(2)	(3)	(4)	(5)	(6)
Adoption of AI	-7.326***	-11.42***	-7.523***	-7.160***	-11.12***	-7.371***
	(2.271)	(2.809)	(2.218)	(2.191)	(2.720)	(2.139)
Fixed Effects:						
– Region	✓	✓	✓	✓	✓	✓
– Industry	✓		✓	✓		✓
– Firm Decile		✓	✓		✓	✓
First Stage F-Stat	47.79	46.52	48.96	46.57	45.25	47.56
Observations	20,057	20,057	20,057	20,064	20,064	20,064

Notes: Standard errors in parentheses. * p <0.10, ** p <0.05, *** p <0.01. Standard errors clustered at the firm level. The independent variable is a dummy, which equals 1 for establishments that did not post AI vacancies in the baseline period, but posted AI vacancies in the endline period. The dependent variables are the change in the log of the respective establishment-level outcomes, specifically the number of non-AI vacancies and the total number of vacancies. AI adoption is instrumented by establishment AI exposure. Each coefficient therefore represents the percentage point impact upon the outcome variable of a one percent rise in the predicted probability of adopting AI. AI exposure is calculated as the standardized average of occupation AI exposure (from Webb 2020), over the occupations for which the establishment posts vacancies in 2010-12, weighted by the number of vacancies posted per occupation, as in Acemoglu et al. (2022).

Table C.B.35: Second stage: Impact of AI adoption on establishment non-AI wages, excluding last year of data

	Growth in Non-AI Median Wage			Growth in Overall Median Wage		
	(1)	(2)	(3)	(4)	(5)	(6)
Adoption of AI	-4.862***	-5.476***	-4.779***	-4.688***	-5.283***	-4.599***
	(1.380)	(1.499)	(1.334)	(1.338)	(1.457)	(1.294)
Fixed Effects:						
– Region	✓	✓	✓	✓	✓	✓
– Industry	✓		✓	✓		✓
– Firm Decile		✓	✓		✓	✓
First Stage F-Stat	47.98	46.62	49.1	46.76	45.35	47.71
Observations	19,888	19,888	19,888	19,895	19,895	19,895

Notes: Standard errors in parentheses. * p <0.10, ** p <0.05, *** p <0.01. Standard errors clustered at the firm level. The independent variable is a dummy, which equals 1 for establishments that did not post AI vacancies in the baseline period, but posted AI vacancies in the endline period. The dependent variables are the change in the log of the respective establishment-level outcomes, specifically wages of non-AI vacancies and wages of all vacancies. AI adoption is instrumented by establishment AI exposure. Each coefficient therefore represents the percentage point impact upon the outcome variable of a one percent rise in the predicted probability of adopting AI. AI exposure is calculated as the standardized average of occupation AI exposure (from Webb 2020), over the occupations for which the establishment posts vacancies in 2010-12, weighted by the number of vacancies posted per occupation, as in Acemoglu et al. (2022).

Table C.B.36: Second stage: Impact of AI adoption on establishment non-AI vacancies, weighted (top 5% winsorized)

	Growth in Non-AI Vacancies			Growth in Total Vacancies		
	(1)	(2)	(3)	(4)	(5)	(6)
Adoption of AI	-2.907**	-4.142***	-2.894**	-2.846**	-4.077***	-2.839**
	(1.298)	(1.437)	(1.264)	(1.279)	(1.419)	(1.246)
Fixed Effects:						
– Region	✓	✓	✓	✓	✓	✓
– Industry	✓		✓	✓		✓
– Firm Decile		✓	✓		✓	✓
First Stage F-Stat	49.54	48.42	51.54	50.7	49.31	52.73
Observations	22,244	22,244	22,244	22,251	22,251	22,251

Notes: Standard errors in parentheses. * p <0.10, ** p <0.05, *** p <0.01. Standard errors clustered at the firm level. Establishments are weighted by baseline number of posts, with the top 5% winsorized. The independent variable is a dummy, which equals 1 for establishments that did not post AI vacancies in the baseline period, but posted AI vacancies in the endline period. The dependent variables are the change in the log of the respective establishment-level outcomes, specifically the number of non-AI vacancies and the total number of vacancies. AI adoption is instrumented by establishment AI exposure. Each coefficient therefore represents the percentage point impact upon the outcome variable of a one percent rise in the predicted probability of adopting AI. AI exposure is calculated as the standardized average of occupation AI exposure (from Webb 2020), over the occupations for which the establishment posts vacancies in 2010-12, weighted by the number of vacancies posted per occupation, as in Acemoglu et al. (2022).

Table C.B.37: Second stage: Impact of AI adoption on establishment non-AI wages, weighted (top 5% winsorized)

	Growth in Non-AI Median Wage			Growth in Overall Median Wage		
	(1)	(2)	(3)	(4)	(5)	(6)
Adoption of AI	-1.135**	-1.382***	-1.088**	-1.111**	-1.364***	-1.064**
	(0.491)	(0.481)	(0.474)	(0.484)	(0.475)	(0.466)
Fixed Effects:						
– Region	✓	✓	✓	✓	✓	✓
– Industry	✓		✓	✓		✓
– Firm Decile		✓	✓		✓	✓
First Stage F-Stat	49.58	48.36	51.54	50.74	49.26	52.74
Observations	22,064	22,064	22,064	22,071	22,071	22,071

Notes: Standard errors in parentheses. * p <0.10, ** p <0.05, *** p <0.01. Standard errors clustered at the firm level. Establishments are weighted by baseline number of posts, with the top 5% winsorized. The independent variable is a dummy, which equals 1 for establishments that did not post AI vacancies in the baseline period, but posted AI vacancies in the endline period. The dependent variables are the change in the log of the respective establishment-level outcomes, specifically wages of non-AI vacancies and wages of all vacancies. AI adoption is instrumented by establishment AI exposure. Each coefficient therefore represents the percentage point impact upon the outcome variable of a one percent rise in the predicted probability of adopting AI. AI exposure is calculated as the standardized average of occupation AI exposure (from Webb 2020), over the occupations for which the establishment posts vacancies in 2010-12, weighted by the number of vacancies posted per occupation, as in Acemoglu et al. (2022).

Table C.B.38: Second stage: Impact of AI adoption on establishment non-AI vacancies, weighting vacancies to match labour survey occupation shares

	Growth in Non-AI Vacancies			Growth in Total Vacancies		
	(1)	(2)	(3)	(4)	(5)	(6)
Adoption of AI	-7.988***	-12.91***	-8.077***	-7.737***	-12.47***	-7.840***
	(2.351)	(3.094)	(2.284)	(2.245)	(2.959)	(2.181)
Fixed Effects:						
– Region	✓	✓	✓	✓	✓	✓
– Industry	✓		✓	✓		✓
– Firm Decile		✓	✓		✓	✓
First Stage F-Stat	43.7	41.58	45.43	44.06	41.83	45.62
Observations	22,244	22,244	22,244	22,251	22,251	22,251

Notes: Standard errors in parentheses. * p <0.10, ** p <0.05, *** p <0.01. Standard errors clustered at the firm level. The independent variable is a dummy, which equals 1 for establishments that did not post AI vacancies in the baseline period, but posted AI vacancies in the endline period. The dependent variables are the change in the log of the respective establishment-level outcomes, specifically the number of non-AI vacancies and the total number of vacancies. AI adoption is instrumented by establishment AI exposure. Each coefficient therefore represents the percentage point impact upon the outcome variable of a one percent rise in the predicted probability of adopting AI. AI exposure is calculated as the standardized average of occupation AI exposure (from Webb 2020), over the occupations for which the establishment posts vacancies in 2010-12, weighted by the number of vacancies posted per occupation, as in Acemoglu et al. (2022). When constructing the establishment-level data set, we weight NIC08 2 digit occupations in our data such that they match the distributions of occupations in the NSS and PLFS.

Table C.B.39: Second stage: Impact of AI adoption on establishment non-AI wages, weighting vacancies to match labour survey occupation shares

	Growth in Non-AI Median Wage			Growth in Overall Median Wage		
	(1)	(2)	(3)	(4)	(5)	(6)
Adoption of AI	-5.181***	-5.733***	-4.977***	-4.933***	-5.472***	-4.735***
	(1.473)	(1.607)	(1.413)	(1.410)	(1.540)	(1.353)
Fixed Effects:						
– Region	✓	✓	✓	✓	✓	✓
– Industry	✓		✓	✓		✓
– Firm Decile		✓	✓		✓	✓
First Stage F-Stat	45.14	42.91	46.86	45.46	43.12	47.02
Observations	22,064	22,064	22,064	22,071	22,071	22,071

Notes: Standard errors in parentheses. * p <0.10, ** p <0.05, *** p <0.01. Standard errors clustered at the firm level. The independent variable is a dummy, which equals 1 for establishments that did not post AI vacancies in the baseline period, but posted AI vacancies in the endline period. The dependent variables are the change in the log of the respective establishment-level outcomes, specifically wages of non-AI vacancies and wages of all vacancies. AI adoption is instrumented by establishment AI exposure. Each coefficient therefore represents the percentage point impact upon the outcome variable of a one percent rise in the predicted probability of adopting AI. AI exposure is calculated as the standardized average of occupation AI exposure (from Webb 2020), over the occupations for which the establishment posts vacancies in 2010-12, weighted by the number of vacancies posted per occupation, as in Acemoglu et al. (2022). When constructing the establishment-level data set, we weight NIC08 2 digit occupations in our data such that they match the distributions of occupations in the NSS and PLFS.

Table C.B.40: Second stage: Impact of AI adoption on establishment non-AI vacancies, keeping only occupations with vacancy shares similar to labour survey occupation shares

	Growth in Non-AI Vacancies			Growth in Total Vacancies		
	(1)	(2)	(3)	(4)	(5)	(6)
Adoption of AI	-7.785*	-13.25***	-8.760*	-7.675*	-13.13***	-8.654*
	(4.523)	(4.959)	(4.497)	(4.476)	(4.937)	(4.459)
Fixed Effects:						
– Region	✓	✓	✓	✓	✓	✓
– Industry	✓		✓	✓		✓
– Firm Decile		✓	✓		✓	✓
First Stage F-Stat	23.7	29.76	24.8	23.45	29.18	24.48
Observations	14,795	14,795	14,795	14,798	14,798	14,798

Notes: Standard errors in parentheses. * p <0.10, ** p <0.05, *** p <0.01. Standard errors clustered at the firm level. The independent variable is a dummy, which equals 1 for establishments that did not post AI vacancies in the baseline period, but posted AI vacancies in the endline period. The dependent variables are the change in the log of the respective establishment-level outcomes, specifically the number of non-AI vacancies and the total number of vacancies. AI adoption is instrumented by establishment AI exposure. Each coefficient therefore represents the percentage point impact upon the outcome variable of a one percent rise in the predicted probability of adopting AI. AI exposure is calculated as the standardized average of occupation AI exposure (from Webb 2020), over the occupations for which the establishment posts vacancies in 2010-12, weighted by the number of vacancies posted per occupation, as in Acemoglu et al. (2022). When constructing the establishment-level data set, we keep NIC08 2 digit occupations whose vacancy shares are similar to these occupations' shares in the NSS and PLFS. In particular, we drop general managers, other professionals, and physical and engingeering science professionals, which appear in our data more than twice as often; and teaching professionals, physical and engineering science associate professionals, teaching associate professionals, customer services clerks, office clerks, models, sales persons and demonstrators, and personal and protective services workers, which appears more than twice as often in labour surveys. This leaves 37% of vacancies with sufficiently similarly distributed occupation shares.

Table C.B.41: Second stage: Impact of AI adoption on establishment non-AI wages, keeping only occupations with vacancy shares similar to labour survey occupation shares

	Growth in Non-AI Median Wage			Growth in Overall Median Wage		
	(1)	(2)	(3)	(4)	(5)	(6)
Adoption of AI	-5.864**	-5.809**	-5.578**	-5.813**	-5.789**	-5.542**
	(2.764)	(2.511)	(2.652)	(2.741)	(2.503)	(2.636)
Fixed Effects:						
– Region	✓	✓	✓	✓	✓	✓
– Industry	✓		✓	✓		✓
– Firm Decile		✓	✓		✓	✓
First Stage F-Stat	23.75	29.8	24.84	23.49	29.22	24.53
Observations	14,683	14,683	14,683	14,686	14,686	14,686

Notes: Standard errors in parentheses. * p <0.10, ** p <0.05, *** p <0.01. Standard errors clustered at the firm level. The independent variable is a dummy, which equals 1 for establishments that did not post AI vacancies in the baseline period, but posted AI vacancies in the endline period. The dependent variables are the change in the log of the respective establishment-level outcomes, specifically wages of non-AI vacancies and wages of all vacancies. AI adoption is instrumented by establishment AI exposure. Each coefficient therefore represents the percentage point impact upon the outcome variable of a one percent rise in the predicted probability of adopting AI. AI exposure is calculated as the standardized average of occupation AI exposure (from Webb 2020), over the occupations for which the establishment posts vacancies in 2010-12, weighted by the number of vacancies posted per occupation, as in Acemoglu et al. (2022). When constructing the establishment-level data set, we keep NIC08 2 digit occupations whose vacancy shares are similar to these occupations' shares in the NSS and PLFS. In particular, we drop general managers, other professionals, and physical and engingeering science professionals, which appear in our data more than twice as often; and teaching professionals, physical and engineering science associate professionals, teaching associate professionals, customer services clerks, office clerks, models, sales persons and demonstrators, and personal and protective services workers, which appears more than twice as often in labour surveys. This leaves 37% of vacancies with sufficiently similarly distributed occupation shares.

Table C.B.42: Second stage: Impact of AI adoption on establishment non-AI vacancies, two-step estimation method

	Growth in Non-AI Vacancies			Growth in Total Vacancies		
	(1)	(2)	(3)	(4)	(5)	(6)
main						
Adoption of AI	-2.567***	-5.112***	-3.072***	-2.728***	-5.108***	-3.258***
	(0.831)	(0.929)	(0.830)	(0.790)	(0.890)	(0.801)
Fixed Effects:						
– Region	✓	✓	✓	✓	✓	✓
– Industry	✓		✓	✓		✓
– Firm Decile		✓	✓		✓	✓
First Stage F-Stat						
Observations	22,392	22,392	22,392	22,399	22,399	22,399

Notes: Standard errors in parentheses. * p <0.10, ** p <0.05, *** p <0.01. The independent variable is a dummy, which equals 1 for establishments that did not post AI vacancies in the baseline period, but posted AI vacancies in the endline period. The dependent variables are the change in the log of the respective establishment-level outcomes, specifically the number of non-AI vacancies and the total number of vacancies. AI adoption is instrumented by establishment AI exposure. Each coefficient therefore represents the percentage point impact upon the outcome variable of a one percent rise in the predicted probability of adopting AI. AI exposure is calculated as the standardized average of occupation AI exposure (from Webb 2020), over the occupations for which the establishment posts vacancies in 2010-12, weighted by the number of vacancies posted per occupation, as in Acemoglu et al. (2022). We use the two-step estimation method implemented using etregress in Stata.

Table C.B.43: Second stage: Impact of AI adoption on establishment non-AI wages, two-step estimation method

	Growth in Non-AI Median Wage			Growth in Overall Median Wage		
	(1)	(2)	(3)	(4)	(5)	(6)
main						
Adoption of AI	-3.381***	-3.540***	-3.363***	-3.211***	-3.353***	-3.183***
	(0.520)	(0.523)	(0.518)	(0.489)	(0.492)	(0.487)
Fixed Effects:						
– Region	✓	✓	✓	✓	✓	✓
– Industry	✓		✓	✓		✓
– Firm Decile		✓	✓		✓	✓
First Stage F-Stat						
Observations	22,214	22,214	22,214	22,221	22,221	22,221

Notes: Standard errors in parentheses. * p <0.10, ** p <0.05, *** p <0.01. The independent variable is a dummy, which equals 1 for establishments that did not post AI vacancies in the baseline period, but posted AI vacancies in the endline period. The dependent variables are the change in the log of the respective establishment-level outcomes, specifically wages of non-AI vacancies and wages of all vacancies. AI adoption is instrumented by establishment AI exposure. Each coefficient therefore represents the percentage point impact upon the outcome variable of a one percent rise in the predicted probability of adopting AI. AI exposure is calculated as the standardized average of occupation AI exposure (from Webb 2020), over the occupations for which the establishment posts vacancies in 2010-12, weighted by the number of vacancies posted per occupation, as in Acemoglu et al. (2022). We use the two-step estimation method implemented using etregress in Stata.

Table C.B.44: Second stage: Impact of AI adoption on establishment non-AI mean wages

	Growth in Non-AI Mean Wage			Growth in Overall Mean Wage		
	(1)	(2)	(3)	(4)	(5)	(6)
Adoption of AI	-5.703***	-6.626***	-5.530***	-5.450***	-6.344***	-5.281***
	(1.403)	(1.580)	(1.345)	(1.343)	(1.516)	(1.289)
Fixed Effects:						
– Region	✓	✓	✓	✓	✓	✓
– Industry	✓		✓	✓		✓
– Firm Decile		✓	✓		✓	✓
First Stage F-Stat	45.14	42.91	46.86	45.46	43.12	47.02
Observations	22,064	22,064	22,064	22,071	22,071	22,071

Notes: Standard errors in parentheses. * p <0.10, ** p <0.05, *** p <0.01. Standard errors clustered at the firm level. The independent variable is a dummy, which equals 1 for establishments that did not post AI vacancies in the baseline period, but posted AI vacancies in the endline period. The dependent variables are the change in the log of the respective establishment-level outcomes, specifically wages of non-AI vacancies and wages of all vacancies. AI adoption is instrumented by establishment AI exposure. Each coefficient therefore represents the percentage point impact upon the outcome variable of a one percent rise in the predicted probability of adopting AI. AI exposure is calculated as the standardized average of occupation AI exposure (from Webb 2020), over the occupations for which the establishment posts vacancies in 2010-12, weighted by the number of vacancies posted per occupation, as in Acemoglu et al. (2022).

Table C.B.45: OLS: Regression of AI adoption on establishment non-AI vacancies

	Growth in Non-AI Vacancies			Growth in Total Vacancies		
	(1)	(2)	(3)	(4)	(5)	(6)
Adoption of AI	1.021***	1.253***	1.160***	1.053***	1.283***	1.189***
	(0.0882)	(0.0852)	(0.0856)	(0.0874)	(0.0842)	(0.0848)
Fixed Effects:						
– Region	✓	✓	✓	✓	✓	✓
– Industry	✓		✓	✓		✓
– Firm Decile		✓	✓		✓	✓
First Stage F-Stat						
Observations	24,312	24,312	24,312	24,322	24,322	24,322

Notes: Standard errors in parentheses. * p <0.10, ** p <0.05, *** p <0.01. Standard errors clustered at the firm level. The independent variable is a dummy, which equals 1 for establishments that did not post AI vacancies in the baseline period, but posted AI vacancies in the endline period. The dependent variables are the change in the log of the respective establishment-level outcomes, specifically the number of non-AI vacancies and the total number of vacancies. AI adoption is instrumented by establishment AI exposure.

Table C.B.46: Probit regression on establishment AI adoption

	AI adoption
Firm Size Decile	-0.0675***
	(0.0121)
Firm Age	0.0367***
	(0.00647)
Total Posts	0.259***
	(0.0153)
Postgraduate Share	9.895***
	(3.067)
Median Salary	-10.75*
	(5.775)
Median Experience	-0.556***
	(0.0624)
Growth in Median Salary	-0.100***
	(0.0285)
Growth in Median Experience	0.457***
	(0.141)
90th Percentile of Salary	0.118**
	(0.0589)
90th Percentile of Experience	-0.308***
	(0.0999)
Growth of 90th Percentile of Salary	0.0591
	(0.0394)
Growth of 90th Percentile of Experience	0.0819
	(0.0741)
99th Percentile of Salary	0.429***
	(0.0497)
99th Percentile of Experience	-2.024***
	(0.626)
Growth of 99th Percentile of Salary	-0.188***
	(0.0330)
Growth of 99th Percentile of Experience	-0.159**
	(0.0639)
Salary Dispersion	-0.00000106***
	(0.000000231)
Experience Dispersion	0.170***
	(0.0493)
N	111044

Notes: Results of a probit regression to compute propensity scores when matching AI adopters to never adopters as described in the short-term results section. All independent variables are lagged by one year. Included but not displayed is a set of year dummies for all years except 2019, and the following interactions: Square of Postgraduate Share, Square and Cube of Median Salary, Square of Median Salary Growth, Square of Growth of 90th Percentile of Salary, Square and Cube of Salary Dispersion, Growth of 99th Percentile of Salary x Salary Dispersion, Median Salary x 99th Percentile of Experience, Growth of Median Experience x 99th Percentile of Experience, Growth of Median Experience x 99th Percentile of Experience x Median Salary. Standard errors clustered at the establishment-level in parentheses. * p <0.10, ** p <0.05, *** p <0.01.

C.B.4 Alternative exposure measures

For our main specifications, we use the AI exposure measure proposed by Webb (2020), which estimates the overlap between occupations' constituent tasks and the capabilities described in AI patents. The advantage of this approach is that it is relatively objective – relying on no 'ad hoc' expert judgements of the capabilities of particular technologies – and flexible, allowing Webb to validate it against previous developments in software and robots. Nonetheless, we also explore alternative AI exposure measures that have been proposed in the literature.

We first consider the AI exposure measure proposed by Felten et al. (2018). Their AI Occupational Impact measure draws on data from the AI Progress Measurement project from the Electronic Frontier Foundation. The data identify nine application areas in which AI has made progress since 2010. Felten et al. (2018) then crowdsource assessments on the applicability of these applications to 52 O*NET ability scales using Amazon MTurk. Their measure assigns an AI exposure score to each O*NET occupation as the weighed sum of the 52 O*NET ability assessments, where the weights are equal to the O*NET-reported prevalence and importance of each ability within each occupation. We map the Felten et al. (2018) measure to Indian NCO using a publicly available crosswalk (see Appendix C.A) and repeat our main specifications. The results are shown in Tables C.B.47 to C.B.49. While we do not observe any significant impacts on the growth of non-AI vacancies for the full sample, we do find similar negative effects on the share of professional and managerial vacancies and on median wage offers.

We also consider the Suitability for Machine Learning (SML) methodology from Brynjolfsson et al. (2018), which uses surveys to score O*NET direct work activities against a rubric of suitability for machine learning (e.g., inputs and outputs are machine-readable, feedback is immediate, task is principally concerned with matching or prediction). We use an India-specific version of the SML index created by Mani et al. (2020), who interviewed a sample of Indian employees using the SML rubric and mapped a SML score onto every occupation in the 2004 NCO at the four-digit level. However, unlike the Webb or Felten et al. measures, we find that the SML exposure measure fails to positively predict establishment AI adoption (Table C.B.50). This may reflect that the SML measure is more forward-looking in its predictions. The result that the SML index does not predict AI demand was also found in Acemoglu et al. (2022), suggesting that the limited predictive power is not limited to India only.

Table C.B.47: Second stage: Impact of AI adoption on establishment non-AI vacancies – Felten et al. exposure measure

	Growth in Non-AI Vacancies			Growth in Total Vacancies		
	(1)	(2)	(3)	(4)	(5)	(6)
Adoption of AI	0.784	2.315	1.508	0.966	2.496	1.750
	(2.541)	(1.682)	(2.569)	(2.669)	(1.739)	(2.709)
Fixed Effects:						
– Region	✓	✓	✓	✓	✓	✓
– Industry	✓		✓	✓		✓
– Firm Decile		✓	✓		✓	✓
First Stage F-Stat	23.91	42.65	22.6	19.53	36.89	18.34
Observations	22,244	22,244	22,244	22,251	22,251	22,251

Notes: Standard errors in parentheses. * p <0.10, ** p <0.05, *** p <0.01. Standard errors clustered at the firm level. The independent variable is a dummy, which equals 1 for establishments that did not post AI vacancies in the baseline period, but posted AI vacancies in the endline period. The dependent variables are the change in the log of the respective establishment-level outcomes, specifically the number of non-AI vacancies and the total number of vacancies. AI adoption is instrumented by establishment AI exposure. Each coefficient therefore represents the percentage point impact upon the outcome variable of a one percent rise in the predicted probability of adopting AI. AI exposure is calculated as the standardized average of occupation AI exposure (from Felten et al. 2018), over the occupations for which the establishment posts vacancies in 2010-12, weighted by the number of vacancies posted per occupation, as in Acemoglu et al. (2022).

Table C.B.48: Second stage: Impact of AI adoption on establishment non-AI vacancy shares, by occupation group – Felten et al. exposure measure

	Growth in Non-AI Vacancies				
	(1)	(2)	(3)	(4)	(5)
Adoption of AI	7.981***	-0.919***	7.608***	-10.31***	-4.039***
	(1.809)	(0.289)	(1.780)	(2.261)	(1.002)
Fixed Effects:					
– Region	✓	✓	✓	✓	✓
– Industry	✓	✓	✓	✓	✓
– Firm Decile	✓	✓	✓	✓	✓
First Stage F-Stat	22.6	22.6	22.6	22.6	22.6
Observations	22,244	22,244	22,244	22,244	22,244

Notes: Standard errors in parentheses. * p <0.10, ** p <0.05, *** p <0.01. Standard errors clustered at the firm level. The independent variable is a dummy, which equals 1 for establishments that did not post AI vacancies in the baseline period, but posted AI vacancies in the endline period. The dependent variables are the change in occupation shares. Occupation groups are 1-digit occupations from the NCO04. Each coefficient therefore represents the impact upon the outcome variable of a one percent rise in the predicted probability of adopting AI. AI exposure is calculated as the standardized average of occupation AI exposure (from Felten et al. 2018), over the occupations for which the establishment posts vacancies in 2010-12, weighted by the number of vacancies posted per occupation, as in Acemoglu et al. (2022).

Table C.B.49: Second stage: Impact of AI adoption on establishment non-AI wages – Felten et al. exposure measure

	Growth in Non-AI Median Wage			Growth in Overall Median Wage		
	(1)	(2)	(3)	(4)	(5)	(6)
Adoption of AI	-3.300**	-2.421**	-3.480**	-3.551**	-2.534**	-3.753**
	(1.507)	(0.958)	(1.560)	(1.662)	(1.020)	(1.729)
Fixed Effects:						
– Region	✓	✓	✓	✓	✓	✓
– Industry	✓		✓	✓		✓
– Firm Decile		✓	✓		✓	✓
First Stage F-Stat	23.71	42.77	22.52	19.35	36.99	18.27
Observations	22,064	22,064	22,064	22,071	22,071	22,071

Notes: Standard errors in parentheses. * p <0.10, ** p <0.05, *** p <0.01. Standard errors clustered at the firm level. The independent variable is a dummy, which equals 1 for establishments that did not post AI vacancies in the baseline period, but posted AI vacancies in the endline period. The dependent variables are the change in the log of the respective establishment-level outcomes, specifically wages of non-AI vacancies and wages of all vacancies. AI adoption is instrumented by establishment AI exposure. Each coefficient therefore represents the percentage point impact upon the outcome variable of a one percent rise in the predicted probability of adopting AI. AI exposure is calculated as the standardized average of occupation AI exposure (from Felten et al. 2018), over the occupations for which the establishment posts vacancies in 2010-12, weighted by the number of vacancies posted per occupation, as in Acemoglu et al. (2022).

Table C.B.50: First stage: Impact of AI exposure on establishment AI adoption – alternative exposure measures

	Adoption of AI					
	(1)	(2)	(3)	(4)	(5)	(6)
AI Exposure	0.00704***	0.0103***	0.00679***	-0.00509***	-0.00922***	-0.00581***
	(0.00159)	(0.00170)	(0.00159)	(0.00144)	(0.00147)	(0.00147)
Exposure Measure	Felten et al.	Felten et al.	Felten et al.	SML	SML	SML
Fixed Effects:						
– Region	✓	✓	✓	✓	✓	✓
– Firm Decile		✓	✓		✓	✓
– Industry	✓		✓	✓		✓
R^2	.0536	.0439	.0607	.0531	.0433	.0604
Observations	22,251	22,251	22,251	22,251	22,251	22,251

Notes: Standard errors in parentheses. * p <0.10, ** p <0.05, *** p <0.01. Standard errors clustered at the firm level. The dependent variable is a dummy, which equals 1 for establishments that did not post AI vacancies in the baseline period, but posted AI vacancies in the endline period. The independent variable is establishment AI exposure, calculated as the standardized average of occupation AI exposure (from either Felten et al. 2018, or Mani et al. 2020 building on Brynjolfsson & Mitchell 2017), over the occupations for which the establishment posts vacancies in 2010-12, weighted by the number of vacancies posted per occupation, as in Acemoglu et al. (2022). Each coefficient therefore represents the impact on the probability of adopting AI of a one-standard deviation rise in AI exposure according to a linear probability model.

Bibliography

Acemoglu, D. (2002), 'Directed technical change', *Review of Economic Studies* **69**(4), 781–809.

Acemoglu, D. (2010), 'When does labor scarcity encourage innovation?', *Journal of Political Economy* **118**(6), 1037–1078.

Acemoglu, D. (2024), 'The Simple Macroeconomics of AI', *NBER Working Paper Series* .

Acemoglu, D., Anderson, G. W., Beede, D. N., Buffington, C., Childress, E. E., Dinlersoz, E., Foster, L. S., Goldschlag, N., Haltiwanger, J., Kroff, Z., Restrepo, P. & Zolas, N. (2024), *Automation and the Workforce: A Firm-Level View from the 2019 Annual Business Survey*, University of Chicago Press, pp. 13–55.

Acemoglu, D. & Autor, D. (2011), Skills, Tasks and Technologies: Implications for Employment and Earnings, *in* 'Handbook of Labor Economics', Vol. 4, Elsevier, pp. 1043–1171.

Acemoglu, D., Autor, D., Hazell, J. & Restrepo, P. (2022), 'Artifial Intelligence and Jobs: Evidence from Online Vacancies', *Journal of Labor Economics* **40**.

Acemoglu, D., Johnson, S. & Robinson, J. (2005), 'The rise of europe: Atlantic trade, institutional change, and economic growth', *American Economic Review* **95**(3), 546–579.

Acemoglu, D. & Restrepo, P. (2019), 'The wrong kind of AI? Artificial intelligence and the future of labour demand', *Cambridge Journal of Regions, Economy and Society* **13**(1), 25–35.

Ackerberg, D. A., Caves, K. & Frazer, G. (2015), 'Identification properties of recent production function estimators', *Econometrica* **83**(6), 2411–2451.

Adams, A., Balgova, M. & Qian, M. (2020), 'Flexible Work Arrangements in Low Wage Jobs: Evidence from Job Vacancy Data', p. 46.

Adão, R., Kolesár, M. & Morales, E. (2019), 'Shift-Share Designs: Theory and Inference*', *The Quarterly Journal of Economics* **134**(4), 1949–2010.

Agarwal, R. (2024), 'The Past and Future of Indian Finance', *India Policy Forum* **20**(1), 193–305. Publisher: National Council of Applied Economic Research.

Agrawal, A., Gans, J. & Goldfarb, A. (2018), *Prediction Machines: The Simple Economics of Artificial Intelligence*, illustrated edition edn, Harvard Business Review Press, Boston, Massachusetts.

Agrawal, A., Gans, J. S. & Goldfarb, A. (2024), 'Artificial intelligence adoption and system-wide change', *Journal of Economics & Management Strategy* **33**(2), 327–337.

Alder, S., Colmer, J. & Coşar, K. (2020), The baltic exchange: Weather risk and international trade in early modern europe.

Aldridge, D. D. (2009), *Admiral Sir John Norris and the British Naval Expeditions to the Baltic Sea 1715–1727*, Nordic Academic Press, Lund.

Alekseeva, L., Azar, J., Gine, M., Samila, S. & Taska, B. (2020), 'The demand for AI skills in the labour market'.

Alfani, G. (2023), 'A step forward toward solving the main mysteries in the history of plague?', *Proceedings of the National Academy of Sciences* **120**(11).

Alfani, G. & Murphy, T. E. (2017), 'Plague and lethal epidemics in the pre-industrial world', *The Journal of Economic History* **77**(1), 314–343.

Alfaro, L. & Chari, A. (2009), 'India transformed? insights form the firm level 1988-2005', *Harvard Business School Working Paper Series* .

Alfaro, L., Fadinger, H., Schymik, J. & Virananda, G. (2024), Industrial and trade policy in supply chains: The case of rare elements.

Allen, B. & Unger, R. (2018), The Allen Unger Commodities Dataset, Technical report.
URL: *https://hdl.handle.net/10622/3SV0BO*

Allen, R. C. (2003), *Farm to Factory: A Reinterpretation of the Soviet Industrial Revolution*, Princeton University Press.

Allen, R. C. (2005), English and welsh agriculture, 1300-1850: Output, inputs, and income.

Allen, R. C. (2009*a*), *The British Industrial Revolution in Global Perspective*, New Approaches to Economic and Social History, Cambridge University Press.

Allen, R. C. (2009*b*), 'Engels' pause: Technical change, capital accumulation, and inequality in the british industrial revolution', *Explorations in Economic History* **46**(4), 418–435.

Allen, T. & Donaldson, D. (2022), 'Persistence and path dependence: A primer', *Regional Science and Urban Economics* **94**, 103724.

Almunia, M., Antràs, P., Lopez-Rodriguez, D. & Morales, E. (2021), 'Venting out: Exports during a domestic slump', *American Economic Review* **111**(11), 3611–3662.

Alonso, C., Berg, A., Kothari, S., Papageorgiou, C. & Rehman, S. (2022), 'Will the AI revolution cause a great divergence?', *Journal of Monetary Economics* **127**(C), 18–37.

Alvarez-Nogal, C., Prados de la Escosura, L. & Santiago-Caballero, C. (2020), 'Economic effects of the black death: Spain in european perspective', *Investigaciones de Historia Económica* **16**(4), 35–48.

Andersen, D. H. (2004), *Mandsmod og kongegunst: En biografi om Peter Wessel Tordenskiold*, Aschehoug, Copenhagen.

Anderson, J. E. & van Wincoop, E. (2003), 'Gravity with Gravitas: A Solution to the Border Puzzle.', *The American Economic Review* **93**, 170–192.

Anderson, J. E. & van Wincoop, E. (2004), 'Trade costs', *Journal of Economic Literature* **42**(3), 691–751.

Andersson, D., Karadja, M. & Prawitz, E. (2020), 'Mass migration and technological change'.

Andersson Palm, L. (2025), 'The database sweden 1570-1810: population, agriculture, land ownership - agricultural statistics on parish level, around 1570, 1630, 1690, 1750, and 1810.'.

Angrist, J. & Pischke, J.-S. (2009), *Mostly Harmless Econometrics: An Empiricist's Companion*.

Antràs, P. & Kulesza, A. (2026), *An "Austrian" Model of International Specialization*.

Antràs, P. & Tubdenov, V. (2025), 'Measuring the average period of production', *AEA Papers and Proceedings* **115**, 624–630.

Arkolakis, C., Ramondo, N., Rodríguez-Clare, A. & Yeaple, S. (2018), 'Innovation and production in the global economy', *American Economic Review* **108**(8), 2128–2173.

Atack, J., Margo, R. A. & Rhode, P. W. (2019), '"automation" of manufacturing in the late nineteenth century: The hand and machine labor study', *Journal of Economic Perspectives* **33**(2), 51–70.

Atkin, D. & Khandelwal, A. K. (2020), 'How distortions alter the impacts of international trade in developing countries', *Annual Review of Economics* **12**(1), 213–238.

Autor, D. H. & Dorn, D. (2013), 'The growth of low-skill service jobs and the polarization of the us labor market', *American Economic Review* **103**(5), 1553–97.

Autor, D. H., Katz, L. F. & Krueger, A. B. (1998), 'Computing inequality: Have computers changed the labor market?', *Quarterly Journal of Economics* **113**(4), 1169–1213.

Autor, D. H., Levy, F. & Murnane, R. J. (2003), 'The Skill Content of Recent Technological Change: An Empirical Exploration', *The Quarterly Journal of Economics* **118**(4), 1279–1333. Publisher: Oxford University Press.

Baasch, E. (1929), 'Zur statistik des ein- und ausfuhrhandels hamburgs anfang des 18. jahrhunderts', *Hansische Geschichtsblätter* **54**, 89–144.

Babina, T., Fedyk, A., He, A. & Hodson, J. (2024), 'Artificial intelligence, firm growth, and product innovation', *Journal of Financial Economics* **151**, 103745.

Babina, T., Fedyk, A., He, A. X. & Hodson, J. (2023), Firm Investments in Artificial Intelligence Technologies and Changes in Workforce Composition, *in* 'Technology, Productivity, and Economic Growth', University of Chicago Press.

Bailey, M., Gupta, A., Hillenbrand, S., Kuchler, T., Richmond, R. & Stroebel, J. (2021), 'International trade and social connectedness', *Journal of International Economics* **129**, 103418.

Bairoch, P., Batou, J. & Chèvre, P. (1988), *La population des villes européennes, 800-1850: banque de données et analyse sommaire des résultats.*

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* pp. 1–14.

Baldwin, R. (2019), *The Globotics Upheaval: Globalisation, Robotics and the Future of Work*, W&N.

Baldwin, R. & Forslid, R. (2020), Globotics and Development: When Manufacturing is Jobless and Services are Tradable, Working Paper 26731, National Bureau of Economic Research.

Baldwin, R. & Taglioni, D. (2007), 'Trade effects of the euro: a comparison of estimators', *Journal of Economic Integration* **22**(4), 780–818.
URL: *http://www.jstor.org/stable/23000919*

Barraclough, G., ed. (1997), *Atlas der Weltgeschichte*, Bechtermünz Verlag, Augsburg.

Baten, J., Szołtysek, M. & Campestrini, M. (2016), '"girl power" in eastern europe? the human capital development of central-eastern and eastern europe in the seventeenth to nineteenth centuries and its determinants', *European Review of Economic History* .

Batra, H., Michaud, A. M. & Mongey, S. (2023), 'Online Job Posts Contain Very Little Wage Information', *Opportunity and Inclusive Growth Institute Working Papers* . Number: 083 Publisher: Federal Reserve Bank of Minneapolis.

Baum-Snow, N. & Han, L. (2024), 'The microgeography of housing supply', *Journal of Political Economy* **132**(6), 1897–1946.

Beestermöller, M. & Rauch, F. (2018), 'A Dissection of Trading Capital: Trade in the Aftermath of the Fall of the Iron Curtain.', *The Journal of Economic History* **78**, 358–393.

bei der Wieden, B. (1999), Die entwicklung der pommerschen bevölkerung, 1701 bis 1918., *in* 'Veröffentlichungen der Historischen Kommission für Pommern (vol. 33). Forschungen zur Pommerschen Geschichte (vol. 5).', Böhlau.

Bengtsson, T., Dribe, M., Quaranta, L. & Svensson, P. (2014), *The Scanian Economic Demographic Database, Version 4.0 (Machine-readable database)*, Lund University, Centre for Economic Demography.

Benhabib, J. & Spiegel, M. M. (2005), Human Capital and Technology Diffusion, Vol. 1 of *Handbook of Economic Growth*, Elsevier, pp. 935–966. ISSN: 1574-0684.

Benzell, S., Lagarda, G. & Rock, D. (2019), Do Labor Demand Shifts Occur Within Firms or Across Them? Non-Routine-Biased Technological Change, 2000-2016, SSRN Scholarly Paper ID 3427396, Social Science Research Network, Rochester, NY.

Bergsøe, A. F. (1853), *Den danske Stats Statistik*.

Berman, E., Bound, J. & Griliches, Z. (1994), 'Changes in the demand for skilled labor within U.S. manufacturing: Evidence from the annual survey of manufactures', *Quarterly Journal of Economics* **109**(2), 367–397.

Bernard, A. B., Moxnes, A. & Saito, Y. U. (2019), 'Production networks, geography, and firm performance', *Journal of Political Economy* **127**(2), 639–688.

Bloom, N., Hassan, T. A., Kalyani, A., Carvalho, M., Lerner, J. & Tahoun, A. (2024), The Diffusion of New Technologies, SSRN Scholarly Paper 3875827, Social Science Research Network, Rochester, NY.

Blūzma, V. (2019), 'Legal regulation of the abolition of serfdom in baltic governorates of the russian empire in the early 19th century: Historical background, realisation, specific features', *Proceedings of the International Scientific Conference: Social Changes in the Global World* **1**(6), 575–589.
URL: *https://js.ugd.edu.mk/index.php/scgw/article/view/3145*

Bohn, R. (1989), *Quellen und Darstellungen zur hansischen Geschichte*, Böhlau, chapter Das Handelshaus Donner in Visby und der gotländische Aussenhandel im 18. Jahrhundert.

Boken, J., Gadenne, L., Nandi, T. & Santamaria, M. (2023), Community networks and trade. CAGE working paper no. 649.

Boon, P. (1993), *Facing the North Sea: West Jutland and the World.*, Esbjerg, chapter The Case of the Westfrisian Seaman Abroad and at Home. Dutch Connections to Scandinavia.

Borusyak, K., Hull, P. & Jaravel, X. (2021), 'Quasi-Experimental Shift-Share Research Designs', *The Review of Economic Studies* **89**(1), 181–213.

Borusyak, K., Jaravel, X. & Spiess, J. (2021), Revisiting Event Study Designs: Robust and Efficient Estimation, Papers 2108.12419, arXiv.org.
URL: *https://ideas.repec.org/p/arx/papers/2108.12419.html*

Bosker, M., Buringh, E. & van Zanden, J. L. (2013), 'From baghdad to london: Unraveling urban development in europe, the middle east, and north africa, 800-1800', *The Review of Economics and Statistics* .

Bosshart, L. & Dittmar, J. (2025), Political competition and economic divergence: Development before and after the black death.

Brancaccio, G., Kalouptsidi, M. & Papageorgiou, T. (2020), 'Geography, transportation, and endogenous trade costs', *Econometrica* **88**(2), 657–691.

Broadberry, S., Campbell, B., Klein, A., Overton, M. & van Leeuwen, B. (2014), *British Economic Growth, 1270–1870*, Cambridge University Press.

Broadberry, S. & Gupta, B. (2009), 'Lancashire, india, and shifting competitive advantage in cotton textiles, 1700-1850: The neglected role of factor prices', *The Economic History Review* **62**(2), 279–305.
URL: *http://www.jstor.org/stable/20542912*

Broadberry, S. N. & Korchmina, E. (2024), Russian economic growth during the eighteenth century.

Brynjolfsson, E., Hui, X. & Liu, M. (2019), 'Does Machine Translation Affect International Trade? Evidence from a Large Digital Platform', *Management Science* **65**(12), 5449–5460.

Brynjolfsson, E., Li, D. & Raymond, L. R. (2023), Generative ai at work, Technical report, National Bureau of Economic Research.

Brynjolfsson, E. & Mitchell, T. (2017), 'What can machine learning do? Workforce implications', *Science* **358**(6370), 1530–1534.

Brynjolfsson, E., Mitchell, T. & Rock, D. (2018), 'What Can Machines Learn, and What Does It Mean for Occupations and the Economy?', *AEA Papers and Proceedings* **108**, 43–47.

Brynjolfsson, E., Rock, D. & Syverson, C. (2017), Artificial Intelligence and the Modern Productivity Paradox: A Clash of Expectations and Statistics, Technical Report w24001, National Bureau of Economic Research.

Burchardi, K. B., Chaney, T. & Hassan, T. A. (2018), 'Migrants, ancestors, and foreign investments', *The Review of Economic Studies* **86**(4), 1448–1486.

Burchardi, K. B. & Hassan, T. A. (2013), 'The economic impact of social ties: Evidence from german reunification', *The Quarterly Journal of Economics* **128**(3), 1219–1271.

Buringh, E. (2021), 'The population of european cities from 700 to 2000: Social and economic history', *Research Data Journal for the Humanities and Social Sciences* **6**(1), 1–18.

Bustos, P. (2011), 'Trade liberalization, exports, and technology upgrading: Evidence on the impact of mercosur on argentinian firms', *American Economic Review* **101**(1), 304–340.

Böhm-Bawerk, E. (1889), *Kapital und Kapitalzins. Zweite Abteilung: Positive Theorie des Kapitales*, Verlag der Wagner'schen Universitäts-Buchhandlung, Innsbruck. Translated into English as *Positive Theory of Capital* by William Smart (London: Macmillan and Co., 1891).

Börschlein, E.-B., Bossler, M. & Popp, M. (2024), Scarce workers, high wages?

Caldara, D., Iacoviello, M. & Yu, D. (2024), Measuring shortages since 1900.

Campante, F. & Yanagizawa-Drott, D. (2017), 'Long-range growth: Economic development in the global network of air links∗′, $TheQuarterlyJournalofEconomics$**133**(3), 1395 − −1458.

Campbell, D. L. (2010), 'History, Culture, and Trade: A Dynamic Gravity Approach', ($\mathrm{EERI}_R P_2 010_2 6$).
URL:

Cantoni, D. (2015), 'The economic effects of the protestant reformation: Testing the weber hypothesis in the german lands', *Journal of the European Economic Association* **13**(4), 561–598.

Carlos, A. M. & Neal, L. (2011), 'Amsterdam and london as financial centers in the eighteenth century', *Financial History Review* **18**(1), 21–46.

Cattaneo, M. D., Crump, R. K., Farrell, M. H. & Feng, Y. (2023), 'On binscatter'.

Cazzaniga, M., Jaumotte, F., Li, L., Melina, G., Panton, A. J., Pizzinelli, C., Rockall, E. J. & Mendes Tavares, M. (2024), 'Gen-AI: Artificial Intelligence and the Future of Work', *Staff Discussion Notes* **2024**(001), 1.

Chahal, H., Abdulla, S., Murdick, J. & Rahkovsky, I. (2021), Mapping India's AI Potential, Technical report, Center for Security and Emerging Technology, Georgetown. Library Catalog: cset.georgetown.edu.

Chen, J. & Roth, J. (2023), 'Logs with Zeros? Some Problems and Solutions*'.

Chor, D. (2010), 'Unpacking sources of comparative advantage: A quantitative approach', *Journal of International Economics* **82**(2), 152–167.
URL: *https://www.sciencedirect.com/science/article/pii/S0022199610000668*

Christiansø Administration (2020), '1709–1720: The Great Northern War', https://www.christiansoe.dk/en/history/1709-1720-the-great-northern-war/.

Clark, G. (2016), 'Microbes and markets: Was the black death an economic revolution?', *Journal of Demographic Economics* **82**(2), 139–165.

Clerides, S. K., Lach, S. & Tybout, J. R. (1998), 'Is learning by exporting important? micro-dynamic evidence from colombia, mexico, and morocco', *The Quarterly Journal of Economics* **113**(3), 903–947.

Cockburn, I. M., Henderson, R. & Stern, S. (2018), The Impact of Artificial Intelligence on Innovation, Technical Report w24449, National Bureau of Economic Research.

Coleman, D. C. (1956), 'Labour in the english economy of the seventeenth century', *The Economic History Review* **8**(3), 280–295.

Conley, T. (1999), 'Gmm estimation with cross sectional dependence', *Journal of Econometrics* **92**(1), 1–45.

Correia, S., Guimarães, P. & Zylkin, T. (2019), 'ppmlhdfe: Fast poisson estimation with high-dimensional fixed effects'.

Costinot, A., Donaldson, D. & Komunjer, I. (2011), 'What goods do countries trade? a quantitative exploration of ricardo's ideas', *The Review of Economic Studies* **79**(2), 581–608.

Cristea, A. D. (2011), 'Buyer-seller relationships in international trade: Evidence from u.s. states' exports and business-class travel', *Journal of International Economics* **84**(2), 207–220.

Davis, D. R. & Weinstein, D. E. (2001), 'An account of global factor trade', *American Economic Review* **91**(5), 1423–1453.

Davis, D. R. & Weinstein, D. E. (2002), 'Bones, bombs, and break points: The geography of economic activity', *American Economic Review* **92**(5), 1269–1289.

de Chaisemartin, C. & D'Haultfoeuille, X. (2022), *Difference-in-Differences Estimators of Intertemporal Treatment Effects*.

De Moor, T. & van Zanden, J. L. (2009), 'Girl power: the european marriage pattern and labour markets in the north sea region in the late medieval and early modern period1', *The Economic History Review* **63**(1), 1–33.

de Vries, J. (2013), *European Urbanization, 1500-1800*, Routledge.

Degn, O. (2017), *The Sound Toll at Elsinore*, Museum Tusculanum Press and the Danish Society for Customs and Tax History.

Degroot, D. (2018), 'Climate change and society in the 15th to 18th centuries', *WIREs Climate Change* **9**(3), e518.
URL: *https://wires.onlinelibrary.wiley.com/doi/abs/10.1002/wcc.518*

Degroot, D., Anchukaitis, K. J., Tierney, J. E., Riede, F., Manica, A., Moesswilde, E. & Gauthier, N. (2022), 'The history of climate and society: a review of the influence of climate change on the human past', *Environmental Research Letters* **17**(10), 103001.

Deming, D. & Kahn, L. B. (2018), 'Skill requirements across firms and labor markets: Evidence from job postings for professionals', *Journal of Labor Economics* **36**(S1), S337–S369.

Denzel, M. A. (2010), *Handbook of World Exchange Rates, 1590–1914*, Ashgate, Farnham.

Di Meo, G. & Eryilmaz, O. (2025), 'The impacts of health shocks on household labor supply and domestic production', *Journal of Health Economics* **101**, 102992.

Dittmar, J. E. & Meisenzahl, R. R. (2019), 'Public goods institutions, human capital, and growth: Evidence from german history', *The Review of Economic Studies* .

Dollinger, P. (1964), *La Hanse.*, Éditions Aubier.

Donaldson, D. & Hornbeck, R. (2016), 'Railroads and american economic growth: A "market access" approach $*', The Quarterly Journal of Economics$**131**(2), $799 - -858$.

Dunlevy, J. A. (2006), 'The Influence of Corruption and Language on the Protrade Effect of Immigrants: Evidence from the American States', *The Review of Economics and Statistics* **88**(1), 182–186.
URL: *https://ideas.repec.org/a/tpr/restat/v88y2006i1p182-186.html*

Eaton, J. & Kortum, S. (2002), 'Technology, Geography, and Trade.', *Econometrica* **70**, 1741–1779.

Edvinsson, R. (2011), 'Shadow interest rates in stockholm and the integration of early financial markets, 1660–1685: was heckscher right?', *Financial History Review* **18**(3), 309–329.

Engström, N. G. (1994), 'Pesten i finland 1710', *Hippokrates. Suomen Lääketieteen Historian Seuran vuosikirja.* **11**, 38–46.

Ericson Wolke, L. (2012), *Sjöslag och rysshärjningar: Kampen om Östersjön under stora nordiska kriget 1700–1721*, Norstedts, Stockholm.

Ewert, U. C. & Selzer, S. (2015), *A Companion to the Hanseatic League*, Brill, Leiden Boston, chapter Social Networks.

Fadinger, H. & Fleiss, P. (2011), 'Trade and sectoral productivity', *The Economic Journal* **121**(555), 958–989.

Fally, T. (2015), 'Structural gravity and fixed effects', *Journal of International Economics* **97**(1), 76–85.
URL: *https://www.sciencedirect.com/science/article/pii/S0022199615001026*

Fang, H., Wang, L. & Yang, Y. (2020), Competition and quality: Evidence from high-speed railways and airlines, Technical report. NBER Working Paper 27475.

Felbermayr, G. J. & Toubal, F. (2010), 'Cultural proximity and trade', *European Economic Review* **54**(2), 279–293.

Felten, E. W., Raj, M. & Seamans, R. (2018), 'A Method to Link Advances in Artificial Intelligence to Occupational Abilities', *AEA Papers and Proceedings* **108**, 54–57.

Fieler, A. C. (2011), 'Nonhomotheticity and bilateral trade: Evidence and a quantitative explanation', *Econometrica* **79**(4), 1069–1101.

Flückiger, M., Hornung, E., Larch, M., Ludwig, M. & Mees, A. (2021), 'Roman transport network connectivity and economic integration', *The Review of Economic Studies* .

Franck, R. (2024), 'Labor scarcity, technology adoption and innovation: evidence from the cholera pandemics in 19th century france', *Journal of Economic Growth* **29**(4), 543–583.

Frandsen, K.-E. (2009), *The Last Plague in the Baltic Region 1709-1713*, Museum Tusculanum Press.

Frost, R. I. (2000), *The Northern Wars: War, State and Society in Northeastern Europe, 1558–1721*, Modern Wars in Perspective, Longman, Harlow.

Gaimster, D. (2005), 'A parallel history: the archaeology of hanseatic urban culture in the baltic c. 1200-1600', *World Archaeology* **37**(3), 408–423.

Gallardo, R. K. & Sauer, J. (2018), 'Adoption of labor-saving technologies in agriculture', *Annual Review of Resource Economics* **10**(1), 185–206.

Galofré-Vilà, G., Hinde, A. & Meera Guntupalli, A. (2018), *Heights across the Last 2,000 Years in England*, Emerald Publishing Limited, pp. 67–98.

Gary, K., Jensen, P. S., Olsson, M., Radu, C. V., Severgnini, B. & Sharp, P. (2022), 'Monopsony power and wages: Evidence from the introduction of serfdom in denmark', *The Economic Journal* **132**(648), 2835–2872.

Gelman, J. R. (1982), 'The english economcy following the black death', *Federal Trade Condition Working Papers* .

Glick, R. & Taylor, A. M. (2010), 'Collateral damage: Trade disruption and the economic impact of war', *The Review of Economics and Statistics* **92**(1), 102–127.

Gmyrek, P., Berg, J. & Bescond, D. (2023), Generative AI and Jobs: A Global Analysis of Potential Effects on Job Quantity and Quality, SSRN Scholarly Paper 4584219, Social Science Research Network, Rochester, NY.

Goldberg, P., Khandelwal, A., Pavcnik, N. & Topalova, P. (2009), 'Trade liberalization and new imported inputs', *American Economic Review* **99**(2), 494–500.

Goldfarb, A., Taska, B. & Teodoridis, F. (2023), 'Could machine learning be a general purpose technology? a comparison of emerging technologies using data from online job postings', *Research Policy* **52**(1), 104653.

Goldsmith-Pinkham, P., Sorkin, I. & Swift, H. (2020), 'Bartik instruments: What, when, why, and how', *American Economic Review* **110**(8), 2586–2624.

Gomtsyan, D. (2022), 'Merchant networks in big cities', *Journal of Urban Economics* **129**, 103440.

Goodman-Bacon, A. (2021), 'Difference-in-differences with variation in treatment timing', *Journal of Econometrics* **225**(2), 254–277.

Goos, M. & Manning, A. (2007), 'Lousy and lovely jobs: The rising polarization of work in britain', *Review of Economics and Statistics* **89**(1), 118–133.

Goos, M., Manning, A. & Salomons, A. (2014), 'Explaining job polarization: Routine-biased technological change and offshoring', *American Economic Review* **104**(8), 2509–2526.

Gould, D. (1994), 'Immigrant links to the home country: Empirical implications for u.s. bilateral trade flows', *The Review of Economics and Statistics* **76**(2), 302–16.
URL: *https://EconPapers.repec.org/RePEc:tpr:restat:v:76:y:1994:i:2:p:302-16*

Greif, A. (1993), 'Contract enforceability and economic institutions in early trade: The maghribi traders' coalition', *The American Economic Review* **83**(3), 525–548.
URL: *http://www.jstor.org/stable/2117532*

Greif, A., Milgrom, P. & Weingast, B. R. (1994), 'Coordination, commitment, and enforcement: The case of the merchant guild', *Journal of Political Economy* **102**(4), 745–776.

Grennan, J. & Michaely, R. (2020), 'Artificial Intelligence and High-Skilled Work: Evidence from Analysts', *SSRN Electronic Journal* .

Griliches, Z. & Mairesse, J. (1984), Productivity and R&D at the firm level, *in* Z. Griliches, ed., 'R&D, Patents, and Productivity', University of Chicago Press, Chicago, pp. 339–374.

Griliches, Z. & Mairesse, J. (1998), Production functions: The search for identification, *in* S. Strom, ed., 'Econometrics and Economic Theory in the Twentieth Century: The Ragnar Frisch Centennial Symposium', Econometric Society Monograph Series, Cambridge University Press, Cambridge.

Groiss, M. & Sondermann, D. (2023), Help wanted: the drivers and implications of labour shortages.

Guinnane, T. W. (2011), 'The historical fertility transition: A guide for economists', *Journal of Economic Literature* **49**(3), 589–614.

Guiso, L., Sapienza, P. & Zingales, L. (2009), 'Cultural biases in economic exchange?', *The Quarterly Journal of Economics* **124**(3), 1095–1131.

Guiso, L., Sapienza, P. & Zingales, L. (2016), 'Long-term persistence', *Journal of the European Economic Association* **14**(6), 1401–1436.

Gøbel, E. (2010), 'The sound toll registers online project, 1497–1857', *International Journal of Maritime History* **22**(2), 305–324.
URL: *https://doi.org/10.1177/084387141002200213*

Habakkuk, H. (1962), *American and British Technology in the Nineteenth Century: The Search for Labour Saving Inventions*, American and British Technology in the Nineteenth Century: The Search for Labour-saving Inventions, Cambridge University Press.
URL: *https://books.google.co.uk/books?id=YJAMAQAAIAAJ*

Hajnal, J. (1965), 'European marriage patterns in perspective', *Population in History: Essays in Historical Demography* pp. 101–143.

Hall, B. H. & Mairesse, J. (1995), 'Exploring the relationship between R&D and productivity in French manufacturing firms', *Journal of Econometrics* **65**(1), 263–293.

Hammel-Kiesow, R. (2000), *Die Hanse.*, C.H. Beck.

Hanson, G. H., Lind, N. & Muendler, M.-A. (2018), The dynamics of comparative advantage.

Harrigan, J., Reshef, A. & Toubal, F. (2020), 'The March of the Techies: Job Polarization Within and Between Firms', *Research Policy* p. 104008.

Hausmann, R. & Rodrik, D. (2003), 'Economic development as self-discovery', *Journal of Development Economics* **72**(2), 603–633.

Head, K. & Mayer, T. (2013), 'What separates us? sources of resistance to globalization', *Canadian Journal of Economics/Revue canadienne d'économique* **46**(4), 1196–1231.

Head, K. & Mayer, T. (2019), 'Misfits in the car industry: Offshore assembly decisions at the variety level', *Journal of the Japanese and International Economies* **52**, 90–105.

Head, K., Mayer, T. & Ries, J. (2010), 'The erosion of colonial trade linkages after independence.', *Journal of International Economics* **81**, 1–14.

Heblich, S., Redding, S. & Zylberberg, Y. (2025), The distributional consequences of trade: Evidence from the grain invasion. CEPR Discussion Paper No. 19522. CEPR Press, Paris London.

Helpman, E. & Krugman, P. R. (1985), *Market Structure and Foreign Trade: Increasing Returns, Imperfect Competition, and the International Economy*, MIT Press, Cambridge, MA.

Hsiang, S. M. (2010), 'Temperatures and cyclones strongly associated with economic production in the caribbean and central america', *Proceedings of the National Academy of Sciences* **107**(35), 15367–15372.

Huang, A., Chilosi, D. & Sapoznik, A. (2019), 'A source collection on urban annuities, 14th–18th centuries'.

Humphries, J. & Schneider, B. (2020), 'Losing the thread: a response to robert allen', *The Economic History Review* **73**(4), 1137–1152.

Humphries, J. & Weisdorf, J. (2019), 'Unreal wages? real income and economic growth in england, 1260–1850', *The Economic Journal* **129**(623), 2867–2887.

International Monetary Fund (2019), 'India: 2019 Article IV Consultation Staff Report', *IMF Staff Country Reports* **19**(385).

International Monetary Fund (2020), Challenges in India's Nonbank Financial Sector, Monetary and Capital Markets Department Special Feature, International Monetary Fund, Washington, D.C.

Jacks, D. S., O'Rourke, K. H. & Taylor, A. M. (2020), 'The Gravitational Constant?', *NBER Working Paper Series* .

Jahnke, C. (2017), *The Sea in History - The Medieval World*, Boydell and Brewer Limited, chapter The Maritime Law of the Baltic Sea, pp. 572–584.

Jarosław Burchardt, J., Meissner, R. K. & Burchardt, D. (2009), 'Bubonic plague – the contagious disease in the great poland and city of poznań area in the first half of the xviii c.', *Nowiny Lekarskie* **78**(1), 79–84.

Javorcik, B., Kett, B., Stapleton, K. & O'Kane, L. (2020), 'Unravelling Trade Integration: Local Labour Market Effects of the Brexit Vote', p. 40.

Jedwab, R., Johnson, N. D. & Koyama, M. (2022), 'The economic impact of the black death', *Journal of Economic Literature* **60**(1), 132–178.

Jedwab, R., Johnson, N. D. & Koyama, M. (2024), 'Pandemics and cities: Evidence from the black death and the long-run', *Journal of Urban Economics* **139**, 103628.

Juhász, R. (2018), 'Temporary protection and technology adoption: Evidence from the napoleonic blockade', *American Economic Review* **108**(11), 3339–3376.

Juhász, R., Lane, N. & Rodrik, D. (2024), 'The new economics of industrial policy', *Annual Review of Economics* **16**(1), 213–242.

Juhász, R., Squicciarini, M. P. & Voigtländer, N. (2024), 'Technology adoption and productivity growth: Evidence from industrialization in france', *Journal of Political Economy* **132**(10), 3215–3259.

Keller, W., Shiue, C. H. & Wang, X. (2021), 'Capital markets in china and britain, 1770–1860: Evidence from grain prices', *American Economic Journal: Applied Economics* **13**(3), 31–64.

Klein, A. & Ogilvie, S. (2015), 'Occupational structure in the czech lands under the second serfdom', *The Economic History Review* **69**(2), 493–521.

Klette, T. J. & Griliches, Z. (1996), 'The inconsistency of common scale estimators when output prices are unobserved and endogenous', *Journal of Applied Econometrics* **11**(4), 343–361.

Klinger, J., Mateos-Garcia, J. & Stathoulopoulos, K. (2018), 'Deep learning, deep change? Mapping the development of the Artificial Intelligence General Purpose Technology', *arXiv:1808.06355 [cs, econ]* .

Klinger, J., Mateos-Garcia, J. & Stathoulopoulos, K. (2020), 'A narrowing of AI research?', *arXiv:2009.10385 [cs]* .

Koch, M., Manuylov, I. & Smolka, M. (2021), 'Robots and Firms', *The Economic Journal* **131**(638), 2553–2584.

Korinek, A. & Stiglitz, J. E. (2021), 'Artificial Intelligence, Globalization, and Strategies for Economic Development', p. 41.

Kossert, A. (2005), *Ostpreußen. Geschichte und Mythos.*, Siedler, chapter Preußische Toleranz, Fremde und 'Repeuplirung', pp. 104–110.

Krizhevsky, A., Sutskever, I. & Hinton, G. E. (2012), ImageNet classification with deep convolutional neural networks, *in* 'Advances in neural information processing systems', Vol. 25.

Kroll, S. (2006), *Städtesystem und Urbanisierung im Ostseeraum in der Frühen Neuzeit: Urbane Lebensräume und Historische Informationssysteme, Beiträge des wissenschaftlichen Kolloquiums in Rostock vom 15. und 16. November 2004. Geschichte und Wissenschaft.*, Stefan Kroll; Kersten Krüger (eds.), chapter Die Pest im Ostseeraum zu Beginn des 18. Jahrhunderts: Stand und Perspektiven der Forschung.

Kroll, S. & Grabinsky, A. (2007), 'Städtesystem und urbanisierung im ostseeraum in der neuzeit – historisches informationssystem und analyse von demografie, wirtschaft und baukultur im 17. und 18. jahrhundert. b: Komplexe historische informationssysteme. b2: Der letzte ausbruch der pest im ostseeraum zu beginn des 18. jahrhunderts. chronologie des seuchenzugs und bestandsaufnahme überlieferter sterbeziffern. karte', University of Rostock.

Krugman, P. (1987), 'The narrow moving band, the dutch disease, and the competitive consequences of mrs. thatcher', *Journal of Development Economics* **27**(1–2), 41–55.

Krugman, P. (1991), 'Increasing returns and economic geography', *Journal of Political Economy* **99**(3), 483–499.

Kuylenstierna, O. (1899), *Striderna vid Gota älfs mynning åren 1717 och 1719*, Norstedt, Stockholm.

Lane, N. (2025), 'Manufacturing revolutions: Industrial policy and industrialization in south korea', *The Quarterly Journal of Economics* .

Langdon, J. (1982), 'The economics of horses and oxen in medieval England', *The Agricultural History Review* **30**(1), 31–40.
URL: *https://www.jstor.org/stable/40274182*

Langdon, J. (1986), *Horses, Oxen and Technological Innovation: The Use of Draught Animals in English Farming from 1066–1500*, Past and Present Publications, Cambridge University Press, Cambridge.

Lee, D. S., McCrary, J., Moreira, M. J. & Porter, J. (2022), 'Valid t-ratio inference for iv', *American Economic Review* **112**(10), 3260–90.

Levinsohn, J. & Petrin, A. (2003), 'Estimating production functions using inputs to control for unobservables', *Review of Economic Studies* **70**(2), 317–341.

Liu, D., Mishra, A. K. & Ray, D. K. (2020), 'Sensitivity of global major crop yields to climate variables: A non-parametric elasticity analysis', *Science of The Total Environment* **748**, 141431.

Loecker, J. D. (2013), 'Detecting learning by exporting', *American Economic Journal: Microeconomics* **5**(3), 1–21.

Luterbacher, J., Dietrich, D., Xoplaki, E., Grosjean, M. & Wanner, H. (2004), 'European seasonal and annual temperature variability, trends, and extremes since 1500', *Science* **303**(5663), 1499–1503.

Magnusson, L. (2007), *An Economic History of Sweden*, Routledge.

Mani, D., Tomar, S., Madan, N. & Bhatia, A. (2020), The Impact of AI on the Indian Labour Market, Technical report, Srini Raju Centre for IT & the Networked Economy, Indian School of Business.

Marczinek, M. (2024), Labour scarcity and productivity: Insights from the last nordic plague.

Marczinek, M., Maurer, S. & Rauch, F. (2025), 'Networks in trade — evidence from the legacy of the hanseatic league', *Journal of International Economics* **157**, 104102.

Maurer, S. & Rauch, F. (2022), 'Economic geography aspects of the panama canal', *Oxford Economic Papers* **75**(1), 142–162.

Mayer, T., Vicard, V. & Zignago, S. (2019), 'The cost of non-europe, revisited', *Economic Policy* **34**(98), 145–199.

McElheran, K., Li, J. F., Brynjolfsson, E., Kroff, Z., Dinlersoz, E., Foster, L. & Zolas, N. (2024), 'Ai adoption in america: Who, what, and where', *Journal of Economics & Management Strategy* **33**(2), 375–415.

McKinsey Global Institute (2019), AI adoption advances, but foundational barriers remain, Technical report, McKinsey.

Melitz, J. (2008), 'Language and foreign trade', *European Economic Review* **52**(4), 667–699.

Melitz, J. & Toubal, F. (2014), 'Native language, spoken language, translation and trade', *Journal of International Economics* **93**(2), 351–363.

Melitz, M. J. (2003), 'The impact of trade on intra-industry reallocations and aggregate industry productivity', *Econometrica* **71**(6), 1695–1725.

Merkel, G. H. (1800), *Die Letten vorzüglich in Liefland am Ende des philosophischen Jahrhunderts. Ein Beytrag zur Völker- und Menschenkunde.*, Graff, Leipzig.

Michaels, G., Rauch, F. & Redding, S. J. (2012), 'Urbanization and structural transformation', *The Quarterly Journal of Economics* **127**(2), 535–586.

Michaels, G., Rauch, F. & Redding, S. J. (2018), 'Task Specialization in U.S. Cities from 1880 to 2000', *Journal of the European Economic Association* **17**(3), 754–798.

Millward, R. (1982), 'An economic analysis of the organization of serfdom in eastern europe', *The Journal of Economic History* **42**(3), 513–548.
URL: *http://www.jstor.org/stable/2120604*

Morrow, P. M. (2010), 'Ricardian–heckscher–ohlin comparative advantage: Theory and evidence', *Journal of International Economics* **82**(2), 137–151.

Munshi, K. (2014), 'Community networks and the process of development', *Journal of Economic Perspectives* **28**(4), 49–76.

Nitsch, V. & Wolf, N. (2013), 'Tear down this wall: on the persistence of borders in trade', *Canadian Journal of Economics/Revue canadienne d'économique* **46**(1), 154–179.

Noy, S. & Zhang, W. (2023), 'Experimental evidence on the productivity effects of generative artificial intelligence', *Science* **381**(6654), 187–192.

Nunn, N. (2007), 'Relationship-specificity, incomplete contracts, and the pattern of trade', *The Quarterly Journal of Economics* **122**(2), 569–600.

Nunn, N. & Puga, D. (2012), 'Ruggedness: The blessing of bad geography in Africa.', *Review of Economics and Statistics* **94**, 20–36.

Nunn, N. & Wantchekon, L. (2011), 'The slave trade and the origins of mistrust in africa', *American Economic Review* **101**(7), 3221–3252.

OECD (2024), Oecd economic outlook, volume 2024 issue 2: Preliminary version, Technical report.

Olden, A. & Møen, J. (2022), 'The triple difference estimator', *The Econometrics Journal* **25**(3), 531–553.

Olley, G. S. & Pakes, A. (1996), 'The dynamics of productivity in the telecommunications equipment industry', *Econometrica* **64**(6), 1263.

Olsson, M. & Svensson, P. (2008), Explaining agricultural growth: The case of Sweden 1700–1850. Paper presented at the 33rd Annual Meeting of the Social Science History Association, Miami, Florida, 23–26 October 2008.

Olsson, M. & Svensson, P. (2017), Historical Database of Scanian Agriculture, Technical report.

Oxford English Dictionary (2020), *Oxford English Dictionary*, Oxford University Press.

O'Rourke, K. & Williamson, J. (2001), *After Columbus: Explaining the Global Trade Boom 1500-1800*.

Parsons, C. & Vézina, P.-L. (2018), 'Migrant networks and trade: The vietnamese boat people as a natural experiment', *The Economic Journal* **128**(612), F210–F234.

Pascali, L. (2017), 'The wind of change: Maritime technology, trade, and economic development', *American Economic Review* **107**(9), 2821–2854.

Perrault, Shoham, Brynjolfsson, Clark, Etchemendy, Grosz, Lyons, Manyika, Mishra & Niebles (2019), The AI Index 2019 Annual Report, Technical report, AI Index Steering Committee, Human-Centered AI Institute, Stanford University, Stanford, CA.

Persson, T. (2011), *Pesten i Blekinge 1710–1711*, Blekinge museum.

Peters, M. E. (2022), 'Government finance and imposition of serfdom after the black death', *European Review of Economic History* **27**(2), 149–173.

Pfister, U. (2017), 'The timing and pattern of real wage divergence in pre-industrial europe: evidence from germany, c. 1500–1850', *The Economic History Review* **70**(3), 701–729.

Pfister, U. (n.d.), Great divergence, consumer revolutionand the reorganization of textile markets: Evidence from hamburg's import trade, eighteenth century.

Pizzinelli, C., Panton, A., Mendes Tavares, M., Cazzaniga, M. & Li, L. (2023), Labor Market Exposure to AI: Cross-Country Differences and Distributional Implications, SSRN Scholarly Paper 4612697, Social Science Research Network, Rochester, NY.

Poeck, D. W. (2010), *Die Herren der Hanse: Delegierte und Netzwerke*, number 8 *in* 'Kieler Werkstücke: Reihe E, Beiträge zur Sozial- und Wirtschaftsgeschichte', Lang.

Pomeranz, K. (2006), 'The economics of respectability: Rural incomes, instability, and gender roles in late imperial china', *Research on Modern Chinese Women's History* p. 205–241.

Puga, D. & Trefler, D. (2014), 'International trade and institutional change: Medieval venice's response to globalization', *The Quarterly Journal of Economics* **129**(2), 753–821.

Pullan, B. (1992), *Epidemics and ideas: essays on the historical perception of pestilence*, Cambridge University Press, chapter Plague and Perceptions of the Poor in Early Modern Italy., p. 101–23.

Raj, M. & Seamans, R. (2018), *Artificial Intelligence, Labor, Productivity, and the Need for Firm-Level Data*, University of Chicago Press, pp. 553–565.

Raster, T. (2023), Contagious coercion: The effect of plagues on serfdom in the baltics.

Raster, T. (2024), Breaking the ice: The persistent effects of pioneers on trade relationships.

Rauch, J. E. (1999), 'Networks versus markets in international trade', *Journal of International Economics* **48**(1), 7–35.

Rauch, J. E. (2001), 'Business and social networks in international trade', *Journal of Economic Literature* **39**(4), 1177–1203.

Rauch, J. & Trindade, V. (2002), 'Ethnic chinese networks in international trade', *The Review of Economics and Statistics* **84**(1), 116–130.

Redding, S. & Venables, A. J. (2004), 'Economic geography and international inequality', *Journal of International Economics* **62**(1), 53–82.

Rock, D. (2019), Engineering Value: The Returns to Technological Talent and Investments in Artificial Intelligence, SSRN Scholarly Paper ID 3427412, Social Science Research Network, Rochester, NY.

Rodger, N. A. M. (2004), *The Safeguard of the Sea*, Penguin.

Romalis, J. (2004), 'Factor proportions and the structure of commodity trade', *American Economic Review* **94**(1), 67–97.

Roosen, J. & Curtis, D. R. (2018), 'Dangers of noncritical use of historical plague data', *Emerging Infectious Diseases* **24**(1), 103–110.

Rosenbaum, P. R. & Rubin, D. B. (1983), 'The central role of the propensity score in observational studies for causal effects', *Biometrika* **70**(1), 41–55.

Sahm, W. (1905), *Geschichte der Pest in Ostpreußen.*

Santos-Silva, J. M. C. & Tenreyro, S. (2006), 'The Log of Gravity.', *The Review of Economics and Statistics* **88**, 641–658.

Scheftel, M. (1999), *Die Hanse. Lebenswirklichkeit und Mythos.*, Schmidt-Römhild, chapter Künstliche Wasserstraßen: Kanäle., pp. 797–800.

Schmelzing, P. (2019), 'Eight centuries of global real interest rates, r-g, and the 'suprasecular' decline, 1311–2018', *SSRN Electronic Journal* .

Schott, P. K. (2003), 'One size fits all? Heckscher–Ohlin specialization in global production', *American Economic Review* **93**(3), 686–708.

Schwartz, P. (1901), *Schriften des Vereins für die Geschichte der Neumark*, chapter Die letzte Pest in der Neumark 1710, pp. 53–79.

Seppel, M. (2019), 'The semiotics of serfdom: How serfdom was perceived in the swedish conglomerate state, 1561–1806', *Scandinavian Journal of History* **45**(1), 48–70.

Shimer, R. (2012), 'Reassessing the Ins and Outs of Unemployment', *Review of Economic Dynamics* **15**(2), 127–148.

Spolaore, E. & Wacziarg, R. (2009), 'The diffusion of development', *Quarterly Journal of Economics* **124**(2), 469–529.

Spruner, K. & Menke, T. (1880), *Hand-Atlas für die Geschichte des Mittelalters und die neueren Zeit*, Justus Perthes, Gotha.

Stephenson, J. Z. (2018), 'Mistaken wages: the cost of labour in the early modern english economy, a reply to robert c. allen', *The Economic History Review* **72**(2), 755–769.

Tammisto, I. (2020), 'Money and grain for law and order: interaction between the nobility and state authorities in livland in the 17th century', *Scandinavian Journal of History* **46**(4), 455–475.

Temin, P. (1966), 'Labor scarcity and the problem of american industrial efficiency in the 1850's', *The Journal of Economic History* **26**(3), 277–298.
URL: *http://www.jstor.org/stable/2115648*

Thiede, F. (1849), *Chronik der Stadt Stettin. Bearbeitet nach Urkunden und den bewährten historischen Nachrichten*, Müller.

Trefler, D. (1993), 'International factor price differences: Leontief was right!', *Journal of Political Economy* **101**(6), 961–987.

Trefler, D. & Sun, R. (2022), AI, Trade and Creative Destruction: A First Look, NBER Working Papers 29980, National Bureau of Economic Research, Inc.

Ulbricht, O. (2004), *Die leidige Seuche. Pest-Fälle in der Frühen Neuzeit.*

Ullman, M. (2006), *Rysskhärjningarna på Ostkusten sommaren 1719*, Bokförlaget Magnus Ullman, Stockholm.

U.S. Census Bureau (2002–2016), 'Annual survey of manufactures, 2002–2021'. Retrieved via Census Bureau API (https://api.census.gov/data/timeseries/asm/). Accessed March 23, 2026.
URL: *https://www.census.gov/programs-surveys/asm.html*

van der Wee, H. (1977), Monetary, credit and banking systems, *in* E. E. Rich & C. H. Wilson, eds, 'The Cambridge Economic History of Europe, Volume V: The Economic Organisation of Early Modern Europe', Cambridge University Press, Cambridge, pp. 290–392.

Vedel, C. (2024), 'A perfect storm: First-nature geography and economic development'.
URL: *https://arxiv.org/abs/2408.00885*

Veluwenkamp, J. W., Scheltjens, W. & Van der Woude, S. (2021), 'Sound toll registers online.', *TSEG - The Low Countries Journal of Social and Economic History* **18**(1), 147–160.

Voigtländer, N. & Voth, H.-J. (2012), 'The three horsemen of riches: Plague, war, and urbanization in early modern europe', *The Review of Economic Studies* **80**(2), 774–811.

von Schöning, K. (1837), 'Aktenmäßige darstellung, wie ein theil von hinterpommern und die provinz neumark brandenburg, als gebiete eines neutralen fürsten, während des nordischen krieges zweimal den unerlaubten durchmarsch feindlicher truppen erfuhren', *Baltische Studien* **4**(1), 46–106.

Voth, H.-J., Caprettini, B. & Trew, A. (2025), Fighting for growth: Labor scarcity and technological progress during the british industrial revolution.

Vourinen, H. S. (2007), *History of plague epidemics in Finland*, Firenze University Press, chapter Histoire des epidemies de peste en Finlande, pp. 53–56.

Waitz, G. (1870), *Hanserecesse. Die Recesse und andere Akten der Hansetage*.

Waldinger, M. (2022), 'The economic effects of long-term climate change: Evidence from the little ice age', *Journal of Political Economy* .

Webb, M. (2020), 'The Impact of Artificial Intelligence on the Labor Market', *SSRN Electronic Journal* .

Weidner, M. & Zylkin, T. (2021), 'Bias and consistency in three-way gravity models', *Journal of International Economics* **132**, 103513.

Winkle, S. (1983), 'Die pest in hamburg. epidemiologische und ätiologische Überlegungen während und nach der letzten pestepidemie im hamburger raum 1712/13', *Hamburger Ärzteblatt* **2**, 1–14.

Wolf, N. (2009), 'Was Germany Ever United? Evidence from Intra- and International Trade, 1885-1933.', *The Journal of Economic History* **69**, 846–881.

Wooldridge, J. M. (2010), *Econometric Analysis of Cross Section and Panel Data*, The MIT Press.

Wubs-Mrozewicz, J. (2017), 'The late medieval and early modern hanse as an institution of conflict management', *Continuity and Change* **32**(1), 59–84.

Wurpts, B. (2020), *The Value of Network Analysis in Historical Sociology: Economic and Social Relations in Medieval Lübeck.*, The Power of Networks., Routledge.

Wurpts, B., Corcoran, K. E. & Pfaff, S. (2018), 'The diffusion of protestantism in northern europe: Historical embeddedness and complex contagions in the adoption of the reformation', *Social Science History* **42**(2), 213–244.

Xoplaki, E. (2005), 'European spring and autumn temperature variability and change of extremes over the last half millennium', *Geophysical Research Letters* **32**(15).

Xu, C. (2022), Reshaping global trade: The immediate and long-run effects of bank failures. Quarterly Journal of Economics, forthcoming.

Xu, G. (2015), 'How does collective reputation affect hiring? selection and sorting in an online labour market', *Mimeo* .

Yotov, Y., Piermartini, R., Monteiro, J.-A. & Larch, M. (2016), *An Advanced Guide to Trade Policy Analysis: The Structural Gravity Model*, WTO Publications.