

Mechanics of Industrial Revolution - eine kritische Auseinandersetzung

Dirk Kalmbach

FernUniversität in Hagen, Fakultät für Wirtschaftswissenschaft

Seminararbeit

August 2024

Studiengang: Master Wirtschaftsinformatik (Teilzeit)

Matrikelnummer: 6124097

Tel.: +86 130 5229 8090

Email: dirk.kalmbach@studium.fernuni-hagen.de

Adresse: 500 Changde Rd., 1-23D, Shanghai, China

Inhaltsverzeichnis

1 Einleitung	1
1.1 Inhalt und Kernaussagen des Papiers	1
1.2 Motivation	2
1.3 Überblick über die Seminararbeit	2
2 Kritisch-theoretische Analyse des Papiers	3
2.1 Das theoretische Modell	3
2.1.1 Entstehungsgründe für Industrialisierung	3
2.1.2 Mögliche weitere Einflussfaktoren auf Industrialisierung	4
2.1.3 Textilindustrie als geeigneter Indikator für Industrialisierung?	6
2.2 Alternative theoretische Herangehensweise	7
3 Kritisch-emprische Analyse	9
3.1 Explorative Datenanalyse (EDA)	9
3.2 Analyse der Regressionsmodelle	12
3.2.1 Regressionsmodell ohne Ausreißer	12
3.2.2 Analyse des räumlichen Aspekts des Regressionsmodells	13
3.2.3 Sensitivitätsanalyse	14
3.2.4 Logistische Regression	15
4 Kritisch-hermeneutische Analyse	17
4.1 Aufbau des Papiers	17
4.2 Autoren	17
4.3 Entstehungsdauer	18
4.4 Wahl des Journals	18
4.5 Wahl des Regressionsmodells	19
4.6 Ungenaue Datenherkunft	20
5 Fazit	20
5.1 Resümee	20
5.2 Zusammenfassung	22
5.3 Nachtrag	22
Literatur	24

A Anhang	25
A.1 Landkreise in England	25
A.2 Variablenbeschreibung	27
A.3 Fehlende Quelle für Textilindustriedaten	29

Abbildungsverzeichnis

1	Anteil Beschäftigter in der Textilindustrie 1831 und 1851	10
2	Boxplot der Beschäftigten in Textilindustrie 1831 und 1851	12
3	Visualisierung des Geosplines	14
4	The Industrial Revolution	23

Tabellenverzeichnis

1	Veränderung des Anteils der Beschäftigten in der Textilbranche zwischen 1831 und 1851	6
2	Landkreise mit höchster absoluter Beschäftigung im Textilsektor	10
3	Landkreise mit niedrigster absoluter Beschäftigung im Textilsektor . . .	11
4	Regressionsanalyse aus Papier (Tabelle 1)	11
5	Determinants of Textile Employment without Outliers	13
6	Regressionsanalyse aus Papier (Tabelle 1) ohne Ausreisser	13
7	Vergleich der Modelfits zwischen Regression mit und ohne Geodaten . .	15
8	Vergleich der mittleren quadratischen Abweichung (MSE) zwischen Regressionsmodell aus Papier (Table 1) und k-fold Cross-Validation (CV)	16
9	Anteil industrialisierter Landkreise 1831 und 1851	16
10	Publikationsstatistiken	18
11	Bewertung des Papiers	21



(entnommen aus Gray, 1824)

1 Einleitung

Diese Seminararbeit beschäftigt sich in kritischer Weise mit dem Papier „The Mechanics of the Industrial Revolution“ von Morgan Kelly, Joel Mokyr und Cormac Ó Gráda - im weiteren einfacherweise mit *Papier* und *Autoren* bezeichnet - veröffentlicht 2023 im *Journal of Political Economy* (Kelly et al., 2023)¹.

1.1 Inhalt und Kernaussagen des Papiers

Das Papier beschäftigt sich auf 26 Kernseiten mit der Entstehung von Industrialisierung im England des 18. Jahrhunderts. Nach Aussage der Autoren handelt es sich dabei um die erste regressionsanalytische Untersuchung der Entstehung von Industrialisierung. Als Indikator für Industrialisierung wählen die Autoren den Anteil der Beschäftigten in der Textilindustrie in den Jahren 1831 und 1851.²

Hierzu führen die Autoren im wesentlichen zwei Regressionen mit der Textilbeschäftigung als Regressant und verschiedenen Regressoren durch. Beide Regressionen sind im Papier unter Tabelle 1 (S. 77) und Tabelle 2 (S. 78) zu finden und werden so auch hier bezeichnet. Weitere Regressionen, vornehmlich um die Plausibilität ihrer Forschungsfrage und die Reliabilität des gewählten Regressionsmodells zu unterstreichen, wurden gerechnet. Dieses Papier konzentriert sich auf die Regression in Tabelle 1.

Untersuchungseinheiten sind bei allen Regressionen die 41 Landkreise Englands³ hinsichtlich ökonomischer Daten wie Beschäftigung, Löhne, Rohstoffe, etc.

Die Kernaussagen sind:

- ★ Industrialisierung begann in Regionen mit einerseits niedrigem Lohnniveau, andererseits vergleichsweise hoher Dichte an Handwerkern in Metallberufen.
- ★ Das Vorhandensein von Banken, Alphabetismus, sowie die Räumliche Nähe zu Kohlevorkommen hatten nur einen geringen Anteil auf die Industrialisierung.
- ★ Entgegen herkömmlicher Meinung stiegen die Reallöhne während der Industrialisierung im weiter vorangeschrittenem, industrialisierendem Norden und sanken

¹Das Papier wurde 2022 zusammen mit dem Datenmaterial vorveröffentlicht. Die folgende Arbeit analysiert diese und zusätzliche Daten mit dem Statistikprogramm R-Studio.

²Erstaunlicherweise gibt es durchaus ergiebiges Datenmaterial aus dem 17. und 18. Jahrhundert, wobei Daten auf Landkreisebene eher selten sind. Zensusdaten wurden gemäß dem britischen Nationalarchiv (vgl. <https://www.nationalarchives.gov.uk/>) beginnend in 1841 bis 1921 alle 10 Jahre jeweils im 2. Quartal erhoben. Kleinere, weniger detaillierte Zensuse gab es auch früher. Eine weitere ergiebige Quelle für Daten auf Landkreisebene ist Marshall 1833.

³Schottland und Wales werden somit nicht berücksichtigt, was die Aussagefähigkeit m.E. nach einschränkt.

im ehemals prosperierenden Süden Englands.

Im Kapitel 4 werde ich dabei noch ausführlicher auf den Inhalt des Papiers eingehen.

1.2 Motivation

Meine Motivation sich in kritischer Weise dem Papier anzunähern, basierte auf der Verwunderung über die sehr hohen R^2 -Werte der Regressionsmodelle. Werte von 0,7 bis 0,9 sind für gewöhnlich gerade bei regionalen Daten eher selten, v.a. wenn man Landkreise als Untersuchungseinheit verwendet, die sich zwar administrativ, nicht aber für gewöhnlich methodisch, d.h. hinsichtlich der Regressoren klar voneinander abtrennen lassen.⁴

Die Autoren verwenden eine spezielle Form der Regressionsanalyse, bei der Teile der erklärenden Variablen als Funktion in das Modell aufgenommen werden. Die Verwendung eines solchen „räumlichen Splines“ im Regressionsmodell überzeugte mich nicht wirklich und erweckte in mir eher den Eindruck, die Analyse „methodisch aufzuwerten“.

Nicht zuletzt war ich überrascht, dass sich solch ein komplexer Prozess wie die Industrialisierung, die ja wie kaum ein anderes Ereignis die Menschheitsgeschichte verändert hat, auf wenige Einflussfaktoren reduzieren lassen sollte.

Dennoch war ich zwiegespalten: der Volkswirt in mir begrüßte die Einfachheit und Eleganz des Modells, während der Soziologe in mir sich weigerte, einen solch komplexen sozio-ökonomisch-technischen sich über einen Zeitraum entwickelnden Prozess wie die Industrialisierung mit einer vgl. einfachen Regression zu erklären.⁵

Der dann aber letztlich ausschlaggebende Grund für die Entscheidung sich kritisch mit dem Papier zu beschäftigen war eine (zunächst) nicht-auffindbare Quelle bei der Überprüfung des Datenmaterials.⁶

1.3 Überblick über die Seminararbeit

Ich entwickle die kritische Auseinandersetzung mit der Arbeit an drei Dimensionen entlang:

⁴Politisch-administrativ sind Landkreise zwar klar anhand der Landkreisgrenzen unterscheidbar, nicht aber zwangsläufig hinsichtlich der betrachteten wirtschaftlich oder sozialen Variablen. Ob ein Arbeiter z.B. in Landkreis A wohnt, schließt ja nicht aus, dass er in Landkreis B arbeitet und ob eine Fabrik in Landkreis C ihren Sitz hat, bedeutet ja nicht, dass sie nicht von Wasserkanälen in anderen Landkreisen, um Rohstoffe und Waren zu verschiffen, profitiert.

⁵Meine erste akademische Ausbildung war ein Studium in Ökonomie, dem ich dann Jahre später ein Fernstudium in Soziologie an der Fernuni anschloss - dies nur zur Einordnung des eben Gesagten.

⁶Wie sich später herausstellte, existierte die Quelle zwar, die Datenbeschreibung im Papier war aber so irreführend, so dass ich an besagter Stelle, keine Daten finden konnte. Hierzu mehr in Kapitel 4.6.

Kapitel 2 liefert eine theoretische Analyse des Papiers. Leitfrage des Kapitels ist: *erklärt das Papier in geeigneter Weise Industrialisierung?* Man könnte auch sagen, dass hier die **Validität** der Forschungsfrage kritisch hinterfragt wird.

Kapitel 3 liefert eine Datenanalyse des von den Autoren dankenderweise bereitgestellten Datenmaterials. Leitfrage des Kapitels ist: *sind die statistischen Verfahren geeignet, Industrialisierung zu erklären?* Man könnte auch sagen, dass hier die **Reliabilität** der Forschungsfrage kritisch hinterfragt wird.

Kapitel 4 versucht in hermeneutischer Weise Inkonsistenzen in dem Papier nachzugehen. Widersprechen sich Aussagen oder gibt es Ungereimtheiten. Leitfrage des Kapitels ist: *sind die Ausführungen im Papier als Ganzes konsistent?* Man könnte auch sagen, dass hier die **Glaubwürdigkeit** der Autoren und des Papiers untersucht wird.

Einen Schwerpunkt bildet dabei die emprische Auseinandersetzung mit dem beigefügten Datenmaterial und der kritischen Überprüfung der statistischen Auswertungen.

Github Alle hier dargestellten statistischen Analysen wurden mit R durchgeführt und sind auf Github unter <https://github.com/dirkkalmbach/Seminar-Applied-Data-Analysis.git> einsehbar. An ausgewählten Stellen wird darüber hinaus in Fußnoten auf die konkrete Stelle im `r-skript` hingewiesen, um ein Auffinden zu erleichtern und die statistischen Erklärungen in der Seminararbeit möglichst knapp zu halten.

2 Kritisch-theoretische Analyse des Papiers

2.1 Das theoretische Modell

2.1.1 Entstehungsgründe für Industrialisierung

Folgende Gründe werden regelmäßig genannt, als mögliche Erklärung für die Entstehung von Industrialisierung und warum diese zunächst in England begann:⁷

- * Holzmangel und ein Überfluss an geeigneten Kohlevorkommen,
- * Großbritannien verfügte über sog. „wet coal“, d.h. Kohle die nah an der Erdoberfläche lagert. Diese konnte einerseits kostengünstig gefördert werden. Andererseits wurden Minen häufig überflutet und man musste eine Methode finden, das Wasser

⁷Vgl. (Khan-Academy, o.A.)

aus den Minen zu bekommen, was zu einem ersten Einsatzgebiet der Dampfmaschine führte,

- ★ eine im Vergleich zu anderen europäischen Ländern geringer Einfluss der Monarchie auf die Gesellschaft verbunden mit vergleichsweise üppiger staatliche Unterstützung für private Geschäftsprojekte,
- ★ eine starke Marine zum Schutz der Handelsschiffe, um Billige Baumwolle, die von in Nordamerika versklavten Afrikanern produziert wurde,
- ★ Profite aus dem transatlantischen Sklavenhandel als Kapitalquelle für Investitionen und
- ★ ein dem Kapitalismus förderlichen Wertesystem wie es Max Weber in seiner *protestantischen Ethik* beschreibt und einer daraus resultierenden geschäftsinteressierten Elite⁸

2.1.2 Mögliche weitere Einflussfaktoren auf Industrialisierung

Die Autoren sind erklärendermaßen Gegner eines Ansatzes zur Erklärung von Industrialisierung, der sich auf die Bedeutung von Kohle und Logistik konzentriert. Zumindest begründen die Autoren, dass in den frühen Jahren der Industrialisierung⁹ Energie und Transportwege keine große Rolle gespielt haben sollen und sehen dies durch ihre Auswertungen bestätigt (vlg. den geringen Einfluss der Regressionskoeffizienten von Kohle und Wasserwegen in Tabelle 4).

⁸Im übrigen werden als Gründe, warum Industrialisierung nicht in China begann, trotz der kulturellen, technischen Überlegenheit Chinas, der Größe des chinesischen Marktes und dem Vorhandensein eines Wasserkanal-Systems, das alle größeren chinesischen Städte verbindet, über dem aktiv Handel betrieben wurde, und dass schon seit Jahrhunderten vor dem des englischen Kanalsystems existierte, folgende mögliche Erklärungen genannt (vgl. Khan-Academy o.A.):

1. Chinas Kohlevorkommen waren weit entfernt von den industriellen Zentren. Außerdem war Chinas Kohle, die weiter im Boden lagert, schwerer zu fördern als die britische Kohle.
2. Die stark ansteigende chinesische Bevölkerung und der daraus resultierenden Verfügbarkeit billiger Arbeitskräfte, die eine Substitution durch Maschinen weniger notwendig machte.
3. Die Konfuzianische Kultur, die gesellschaftliche Stabilität höher bewertet als Veränderung und Fortschritt.
4. Hohe Verteidigungsausgaben aufgrund der langen West- und Nordgrenze Chinas.
5. Eine zunehmende Abschottung Chinas gegenüber der Welt während der Qing-Dynastie.

⁹Wobei fraglich ist, ob das 18. Jahrhundert als frühe Phase von Industrialisierung angesehen werden kann, wie das nächste Unterkapitel erläutert.

Viele andere Autoren sehen das allerdings anders, was eine allgemeine online Recherche leicht zeigt. Aber auch ohne in die wissenschaftliche Diskussion einzusteigen, kann die Konstruktion der beiden Variablen kritisiert werden: die Variable `Coal_Distance` wird durch die Distanz zwischen Landkreismitte und nächstem Kohlebergwerk gemessen, wobei Landkreise mit Kohleabbau pauschal mit 20 km angesetzt wurden. Letztendlich soll die Variable ein Platzhalter für den Kohlepreis sein. Inwieweit die so gemessene Kohledistanz ein geeigneter Indikator ist kann stark angezweifelt werden. Um besser einordnen zu können, wie sich die Landkreise in England verteilen, liefert die Abbildung auf S. iv einen Überblick. Man sieht dass gerade in Zentralengland viele kleinere und „längliche“ Landkreise mit oftmals nur 20 km Durchmesser existieren. Das wiederum wirft die Frage auf, ob die so modellierte Kohledistanz wirklich geeignet ist, die Verfügbarkeit von Kohle im jeweiligen Landkreis geeignet abzubilden. Das wiederum wäre ein Hinweis darauf, warum die Variable kaum Erklärungsbeitrag im Regressionsmodell liefern kann, was die geringen Regressionskoeffizienten verdeutlichen. Darauf hinaus bleibt ungeklärt, warum die Erklärungskraft der Kohle im Modell im Zeitverlauf abnimmt. Die Begründung der Autoren ist ja, dass Kohle erst später seine große Bedeutung erlangt. Dann aber müsste der Regressionskoeffizient der `Coal_Distance` für 1851 größer sein als für 1831.

Dazu kommt, dass die direkte Distanz zum Kohlebergwerk (Luftlinie) allein wenig Aussagekraft hat. Entscheidend sind vielmehr die Transportwege (Straßen, Flüsse) und die verfügbaren Transportmöglichkeiten im jeweiligen Landkreis und den jeweils angrenzenden Landkreisen. Das wirft die Frage der Bedeutung von Logistik für die Entstehung von Industrialisierung auf.

Die Variable `Water_Flow` wird anhand der Wassermenge, die durch den Landkreis fließt, gemessen. Die Variable misst somit einerseits Energie (Wasserkraft) als auch Transport (Wasserwege). Die große Bedeutung des Schiffsverkehrs (insb. für Landkreise mit Küste wichtig) wird nicht berücksichtigt.

Industrialisierung schon im 16. Jahrhundert? Es gibt Hinweise, dass die Industrialisierung schon weit vor dem 18. Jahrhundert begann: Auswertungen von Arbeitsstatistiken aus dem 17. Jahrhundert - und somit bereits Generationen vor der hier untersuchten Zeit von 1831 und 1851 - deuten auf eine frühe wirtschaftliche Transformation hin: Landwirtschaft nahm zugunsten des produzierendes Gewerbes (Handwerksbetriebe und von zu Hause arbeitende Weber) ab. Darüber hinaus ist interessant, dass sich ebenso bereits im 19. Jahrhundert der Dienstleistungssektor verdoppelte - und nicht erst Mitte des 20. Jahrhunderts. Zumindest in England begannen die wirtschaftlichen

VERÄNDERUNG BESCHÄFTIGUNG IN TEXTILBRANCHE 1831-1851		n
Anzahl Counties mit Wachstum der Beschäftigung		17
Anzahl Counties mit Rückgang der Beschäftigung		24

Tabelle 1: Veränderung des Anteils der Beschäftigten in der Textilbranche zwischen 1831 und 1851

Transformationsphasen also deutlich früher als allgemein angenommen.¹⁰ ¹¹

2.1.3 Textilindustrie als geeigneter Indikator für Industrialisierung?

Ist Industrialisierung überhaupt monokausal. Muss nicht vielmehr angenommen werden, dass es sich um einen komplexen Prozess handelt mit gegenseitig beeinflussenden Variablen, der sich regional und zeitlich entfaltet und sowohl ökonomisch, sozial, politisch und technisch messen lässt?

Wenngleich die Textilindustrie gemeinhin für die Schlüsselindustrie der frühen Industrialisierung gilt, muss man doch fragen, ob es sinnvoll ist, Industrialisierung mit Textilbeschäftigung gleichzusetzen. Zweckmäßiger wäre wohl einen Indikator zu bilden, der mehrere Aspekte der Industrialisierung (Industriebeschäftigte, Bevölkerungswachstum, Urbanisierung, Kapital, etc.) umfasst, wobei dann wiederum die Frage nach Ursache und Wirkung gestellt werden muss.

Anteil der Beschäftigten in der Textilbranche ist rückläufig Ein weiterer Hinweis, dass die relative Höhe des Beschäftigungsanteils in der Textilindustrie vermutlich kein guter Indikator für Industrialisierung ist, zeigt der rückläufige Anteil der Texilbeschäftigen in der Mehrheit der Landkreise (siehe Tabelle 1). Dem könnte man nun entgegen halten, dass die Textilindustrie lediglich den Beginn der Industrialisierung repräsentierte und sozusagen ein industrieller Frühindikator war. Machinelle Produktionsverfahren begannen in der Textilindustrie und weiteten sich auf andere Produkte und Industrien aus.

Das mag einerseits daran liegen, dass eine gewisse Sättigung in der Textilbranche relativ zu anderen Branchen der Industrialisierung eingesetzt hatte. Wie gesagt, handelt es sich um relative Beschäftigungsanteile. Absolut hatte England ein Bevölkerungswachstum zwischen 1831 und 1851 von 13 Mio auf fast 17 Mio Einwohner.¹²

¹⁰Vgl. Uni (2024)

¹¹Aus Zeit- und Platzgründen im Rahmen dieser Seminararbeit zwar nicht möglich, wäre es durchaus interessant zu überprüfen, ob damals schon Daten für einzelne Landkreise erhoben wurden und - falls so - zu überprüfen, ob sich statistische Zusammenhänge zwischen dem Grad der Transformation im 16. Jahrhundert und dem Auftreten von Textilindustrien im 19. Jahrhundert erkennen lassen.

¹²Vgl. die Zensusdaten in <https://www.dittobooks.co.uk/extras/census/census-population-1801->

Schaut man sich darüber hinaus die geografische Verteilung der Counties an, überrascht, dass es v.a. Counties im Süden und Osten sind - jene Regionen also in denen die Industrialisierung gemeinhin später begonnen hat

2.2 Alternative theoretische Herangehensweise

Einleitung Industrialisierung ist nicht direkt messbar. Es handelt sich um ein theoretisches Konstrukt, das zunächst durch geeignete Indikatoren messbar gemacht werden muss. Die Autoren wählen als Indikator für Industrialisierung den Anteil der Beschäftigten der Textilindustrie.

Es wäre also konsequenter zu reklamieren, dass die Autoren die Entstehung der Textilindustrie im beginnenden industriellen England regressionsanalytisch erfasst haben.

Eine Konzentration auf eine einzige Variable wie Textilbeschäftigung halte ich für wenig zielführend. Stattdessen sollte ein Faktor entwickelt werden, der eine Vielzahl an Variablen beinhaltet, an der man die Sichtbarkeit von Industrialisierung abtragen kann, um der Komplexität des Industrialisierungsphänomens Rechnung zu tragen. Idealerweise wäre ein solcher Faktor als Zeitreihe konstruierbar.

Die Vielfalt und Unübersichtlichkeit an möglichen Erklärungen liegt nahe, auf abstraktere Erklärkonzepte zurückzugreifen. Ein mögliches Erklärungsmodell wäre im Rahmen eines systemtheoretischen Evolutionsverständnis Industrialisierung als zunehmende Ausdifferenzierung des Wirtschaftssystem in einer funktional differenzierten Gesellschaft zu sehen.

Evolution Evolution besteht aus Variation, Selektion und Stabilisierung. In biologischen Systemen beruht Variation auf Zufall, in einem sozio-ökonomischen System kann Variation auch mit Innovation gleichgesetzt werden.

"Evolutionäre Veränderungen des Wirtschaftssystems als ganzem [!sic] ergeben sich, so die systemtheoretische Vorstellung, wenn die positive Selektion einer neuen Variation in der Form der kommunikativen Verständigung über die Zuordnung knapper Güter zu einer Systemdifferenzierung führt. Die innere Differenzierung des Wirtschaftssystems ist das Ergebnis der Reintegration des Variations- und Selektionsmechanismus, wenn sich das System in einer den neuen Umweltanforderungen gerechten Struktur stabilisiert hat." (Lierl 2004, S. 11)

Meine These wäre jetzt, Industrialisierung als zeitliche Phase der vermehrt zunehmenden Verselbständigung des Kommunikationsmediums Geld aufzufassen.

Empirische Überprüfbarkeit Diese kommunikationstheoretische Betrachtung von Industrialisierung hätte den Vorteil der Messbarkeit: Textdokumente des 17., 18. und 19. Jahrhunderts sind sowohl für England als auch für andere Länder wie Niederlande, Frankreich, Deutschland, USA etc. umfassend vorhanden und könnten vergleichsweise einfach mittels automatisierter Textauswertung (NLP) auch in großen Mengen analysiert werden.

Da vermutet werden kann, dass allgemein Kommunikation häufig bereits vor sozialer oder ökonomischer Handlung/Ereignisse sichtbar ist, könnte folglich Kommunikation als Frühindikator für Industrialisierung dienen und aufgrund der Verfügbarkeit von Daten (Textdokumente), eine Ländervergleichsanalyse durchgeführt werden. Als Untersuchungseinheit wären somit Länder denkbar, die ja an sich aufgrund der Sprache abgeschlossene Kommunikationsräume darstellen.

Mein Vorschlag wäre somit, eine Analyse von historischen Textdokumenten zur Untersuchung der unterschiedlichen Entwicklung von Industrialisierung in verschiedenen Ländern durchzuführen. Diese könnte in grob Schritten beschrieben folgendermaßen aussehen: nachdem der Untersuchungszeitraum festgelegt wurde (z.B. 1700 bis 1850) würden in einem ersten Schritt Dokumente aus dieser Zeit gesichtet und entsprechende Codes identifiziert, die als Indikator für Gewinnaussichten in Frage kämen. Die im Codierungsschritt ermittelte zeitliche Häufigkeit des Auftretens von Schlüsselbegriffen in Texten wäre dann die unabhängige Variable.

Leitfragen wären dann u.a.:

- ★ Gibt es kausale Zusammenhänge zwischen bestimmten Kommunikationsbegriffen und Industrialisierung?
- ★ Lässt sich Industrialisierung als eine Phase "radikaler" Ausdifferenzierung des Wirtschaftssystem auffassen? Und wenn ja, bringt dies einen Erkenntnisgewinn zur Erklärung von Industrialisierung?
- ★ Warum entstand Industrialisierung zuerst in England?
- ★ Wie unterschieden sich andere Länder hinsichtlich gesellschaftlicher Kommunikation während der entstehenden Industrialisierung?

3 Kritisch-emprische Analyse

3.1 Explorative Datenanalyse (EDA)

Ungenauigkeiten bei der Data Preperation

Im Rahmen der EDA fielen einige Ungenauigkeiten auf. Auffallend waren einige Berechnungsfehler bei den konstruierten Variablen im Datensatz `england_map.Rda`:

- * Die Variablen für die Bevölkerungszahlen für 1761 und 1831 bestehen aus unterschiedlichen Datentypen.¹³ Das hat jetzt zwar keine Auswertung auf das Ergebnis, zeigt aber dass die Autoren das Data Cleaning wohl nicht so genau genommen haben.
- * Bei der Berechnung der Quantile für die Variable `Textiles_1831` wurde ein Landkreis (YNR) falsch zugeordnet.¹⁴ Auch dies hat keine Auswirkungen auf die Regressionsmodelle, verzerrt aber die grafischen Ausdrucke der Landkreise im Papier zugunsten der Interpretation der Autoren.
- * Für die Quantilbildung der Variablen `wage_1760` gaben sich sogar Abweichungen für 5 verschiedene Landkreise. Wenngleich diese gering waren, ändert sich auch hier die Interpretation der grafischen Ergebnisse zugunsten der Autoren. Die Quantilbildung der Variablen `wage_1830` hingegen ist fehlerfrei.
- * Die Variable `Wage` wurde durch Logarithmierung der Rohdatenvariable `wage.1767` gebildet. Hier zeigte sich dass für den Landkreis Herfordshire (HRF) ein falsches Ergebnis mit einer deutlichen Abweichung berechnet wurde.¹⁵

Nun kann man diese kleineren „Ungenauigkeiten“ als Leichtigkeitsfehler abtun, man muss sich aber schon fragen, wie die Autoren mit der Datenanalyse grundsätzlich vorgegangen sind. Solche Abweichungen, die nur einzelne Werte betreffen dürften ja eigentlich nicht vorkommen, da im Data Cleaning und Data Preperation die Schritte nicht nur nachvollziehbar, sondern v.a. Variablenweise erfolgten, d.h. z.B. dass bei einer Logarithmierung einer Variablen ein Wert deutlich anders ist, kann eigentlich nicht vorkommen.

Absolute Beschäftigung in der Textilindustrie *Wie hoch sind eigentlich die absoluten Beschäftigungszahlen in der Textilindustrie in den einzelnen Landkreisen?* Der

¹³Vgl. den Abschnitt „POPULATION“ in `cross-check.R`

¹⁴Vgl. den Abschnitt „CHECK QUANTILES“ in `cross-check.R`

¹⁵Vgl. den Abschnitt „CHECK THE VAR wage.1767 AND WAGE“ in `cross-check.R`

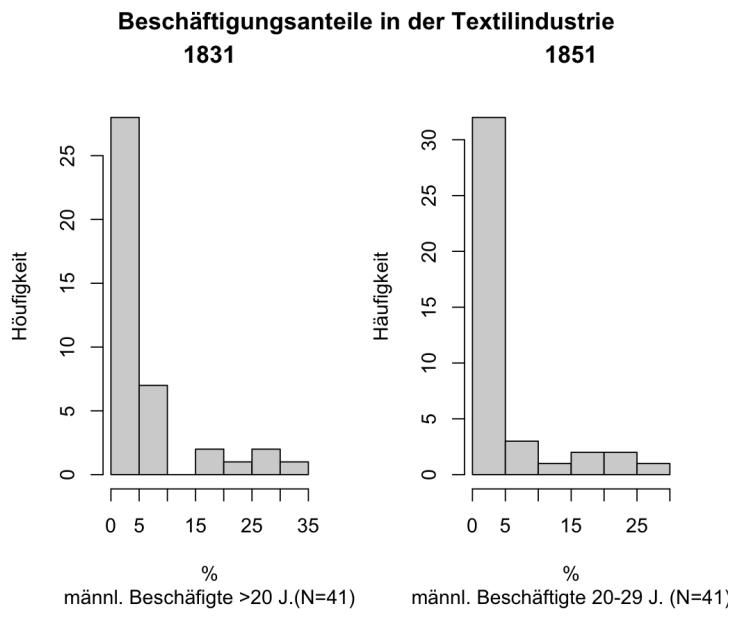


Abbildung 1: Anteil Beschäftigter in der Textilindustrie 1831 und 1851

	Code	Name	Textiles 1831	Textiles 1851
4	CHE	Cheshire	28627	5961
6	CNW	Cornwall	21011	5294
22	MSX	Middlesex	15805	2489
40	YNR	Yorkshire.North	14005	4320
10	DRS	Dorset	10854	2986
9	DRH	Durham	10693	2703

Tabelle 2: Landkreise mit höchster absoluter Beschäftigung im Textilsektor

Datensatz beinhaltet ja logarithmierte Anteilswerte. Dadurch sind die tatsächlichen Größenverhältnisse nicht sofort sichtbar. Eine kurze Datentransformation zeigte in der Tat, dass einige Landkreise so geringe Werte aufweisen, dass man die Anwendung von multivariaten Analyseverfahren wie Regressionsanalysen durchaus in Frage stellen kann. Andere Landkreise wiederum weisen sehr hohe absolute Beschäftigungszahlen auf (vgl. Tabellen 2 und 3). Der Datensatz besteht also aus Extremwerten.

Und in der Tat weisen einige wenn auch wenige Landkreise so geringe Beschäftigungszahlen auf, dass die Frage erlaubt ist, ob diese Landkreise nicht besser aus der Analyse entfernt werden sollten.

Außerdem beruhen die Daten auf Zensuserhebungen. Inwieweit diese Erhebungsdaten damals wirklich so genau waren kann ebenfalls angzweifelt werden: ein kleiner Messfehler in Landkreisen mit niedriger Beschäftigungszahlen könnte durchaus zu einer deutlichen Verzerrung des Regressionsschätzers führen.

	Code	Name	Textiles 1831	Textiles 1851
29	SFF	Suffolk	82	93
38	WTS	Wiltshire	67	201
27	OXD	Oxfordshire	42	38
28	RTL	Rutland	38	14
15	HNT	Huntingdonshire	15	11
35	WML	Westmorland	15	24

Tabelle 3: Landkreise mit niedrigster absoluter Beschäftigung im Textilsektor

	(1)	(2)	(3)	(4)	(5)	(6)
Skills 1790s	2.022*** (0.503)	2.466*** (0.668)	2.020*** (0.511)	1.151*** (0.398)	1.079* (0.532)	1.007** (0.377)
Wage 1760s	-4.776* (2.742)	-5.402* (2.817)	-4.783 (2.937)	-3.035 (2.109)	-3.004 (2.138)	-1.623 (2.056)
Mkt Potential 1750	0.712** (0.301)	0.592* (0.320)	0.712** (0.306)	0.503** (0.217)	0.514** (0.227)	0.403* (0.209)
Distance to Coal		0.720 (0.685)			-0.127 (0.650)	
Water Power			0.001 (0.425)			-0.664* (0.348)
Num.Obs.	41	41	41	41	41	41
R2	0.684	0.695	0.676	0.813	0.810	0.865
R2 Adj.	0.684	0.695	0.676	0.8135	0.810	0.865
AIC	131.1	130.6	132.9	88.7	89.5	74.3
s(X,Y)	0.641	0.534	0.768	0.070	0.076	0.0132

Tabelle 4: Regressionsanalyse aus Papier (Tabelle 1)

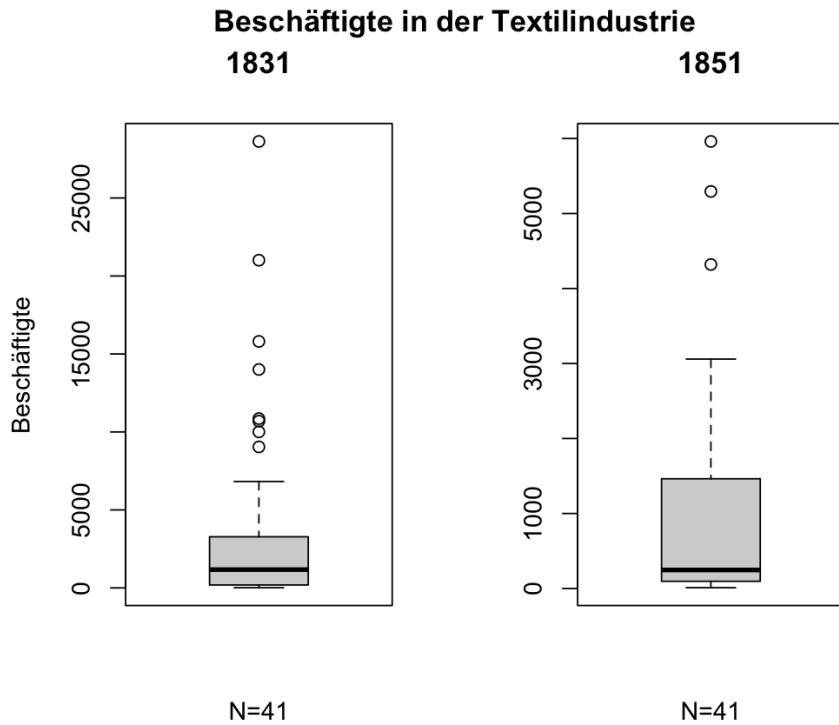


Abbildung 2: Boxplot der Beschäftigten in Textilindustrie 1831 und 1851

3.2 Analyse der Regressionsmodelle

Einige allgemeine Anmerkungen zu den Modellen Auffällig ist auch, dass die Regressionsanalyse für 1851 einen geringeren Regressionskoeffizienten für den Einfluss der Kohle (`Coal_Distance`) liefert als für 1831. Letztendlich müsste der Einfluss der Kohle doch zunehmen und nicht abnehmen.

Was auch in dem Regressionsoutput im Papier fehlt, sind die Daten für die Regressionskonstante. Die Regressionskonstante könnte interpretiert werden als „natürliche“ Erklärung für Industrialisierung, d.h. wie viel Industrialisierung erklärt das Modell, wenn alle unabhängigen Variablen null sind. Allerdings ist diese in allen Modellen nicht signifikant. Sie reicht von 4,1 bis 10,5.¹⁶

3.2.1 Regressionsmodell ohne Ausreißer

Die Boxplots in Abbildung 2 zeigen die Ausreißer in den abhängigen Variablen. Ausreißer sind jene Fälle die in der untersuchten Variablen einen Wert außerhalb des 1,5 fachen Interquartilsabstands aufweisen (sog. extreme Ausreißer). Einer gängigen Praxis in der Datenauswertung folgend, habe ich im folgenden diese Landkreise entfernt (so-

¹⁶Vgl. Abschnitt „Tabelle 1 aus Papier replizieren“ in `cross-check.R`

Tabelle 5: Determinants of Textile Employment without Outliers

	(1)	(2)	(3)	(4)	(5)	(6)
Skills 1790s	2.195*** (0.556)	3.161*** (0.671)	2.207*** (0.570)	1.582** (0.449)	1.475** (0.557)	1.438*** (0.421)
Wage 1760s	-2.989 (3.510)	-4.102*** (0.924)	-3.323 (3.725)	-6.043* (2.876)	-5.969* (2.915)	-4.834 (2.639)
Mkt Potential 1750	0.712** (0.335)	0.418 (0.351)	0.701* (0.344)	0.386 (0.271)	0.399 (0.276)	0.218 (0.266)
Distance to Coal		1.336* (0.687)			-0.201 (0.773)	
Water Power			0.108 (0.481)			-0.753* (0.351)
Num.Obs.	32	32	32	32	32	32
R2	0.667	0.701	0.658	0.821	0.827	0.902
R2 Adj.	0.667	0.701	0.658	0.821	0.827	0.902
AIC	105.0	101.1	106.7	63.7	61.9	39.3
s(X,Y)	0.531	0.021	0.872	0.0554	0.0542	0.0159

Tabelle 6: Regressionsanalyse aus Papier (Tabelle 1) ohne Ausreisser

wohl für die abhängigen als auch unabhängigen Variablen) und die Regressionsanalysen der Autoren wiederholt.¹⁷ Es blieben 32 Landkreise übrig.

Der Regressiounsoutput ist in Abbildung 6 dargestellt. Das Modell zeigt deutlich bessere Modellfit-Werte, während sich die Signifikanz der Regressionskoeffizienten nicht wesentlich ändert.

3.2.2 Analyse des räumlichen Aspekts des Regressionsmodells

Visualisierung des Geosplines Den Versuch einer Visualisierung des räumlichen Splines zeigt Abb. 3. Hierbei wurden die in der Regression aus Tabelle 1 im Papier geschätzten Werte für die Variablen X und Y, also die Koordinaten der einzelnen Landkreise über die Landkarte von England gelegt.¹⁸ Die schwarzen Linien zeigen dabei die Schätzfunktion, die roten und grünen Linien jeweils die untere und obere Standardabweichung ($\pm 1s.e.$). Man bekommt eine ungefähre Vorstellung wie das Modell die Daten

¹⁷Sofern es sich nicht um systematische Ausreisser handelt.

¹⁸X und Y sind die Längen- und Breitengrade, die als Funktion in die Regression aufgenommen worden und deren Schätzwert im Gegensatz zu den Regressionskoeffizienten als Funktion aufgefasst werden müssen.

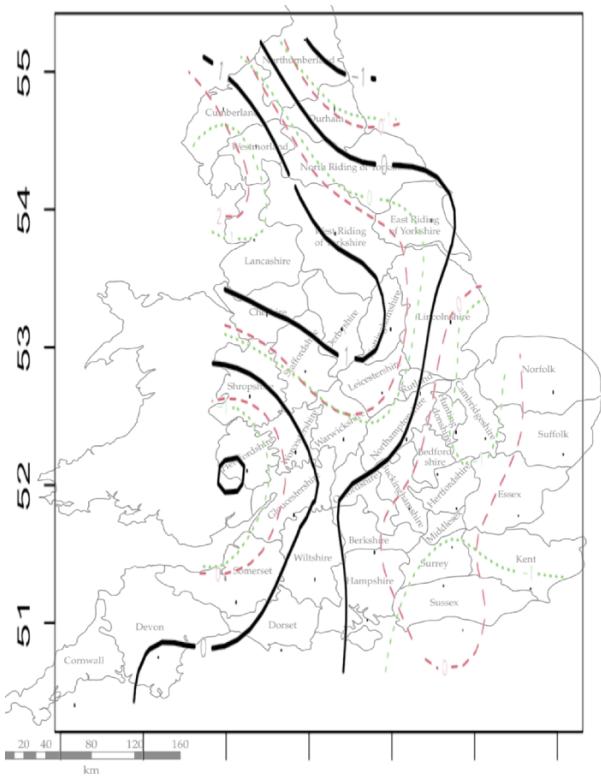


Abbildung 3: Visualisierung des Geosplines

räumlich trennt.¹⁹

Vergleich der Fits Wie „schlägt“ sich eigentlich ein einfaches multiples Regressionsmodell ohne Geodaten mit dem (vergleichsweise schwer zu interpretierenden) additiven Modell? Hierzu habe ich für die sechs Modellvarianten der Regression aus Tabelle 1 im Papier ein Regression gerechnet. Der Vergleich der Model-Fit Werte überrascht wie Tabelle 7 zeigt.²⁰ Wie man sieht, sind die die Werte für die ersten 3 Modelle, also den Modellen, die die Textilbeschäftigung von 1831 prognostizieren, nahezu identisch.

3.2.3 Sensitivitätsanalyse

Ziel einer Sensitivitätsanaylse ist es, zu ermitteln, welchen Einfluss die Varianz der Einflussvariablen auf das Regressionsmodell haben. Eine Möglichkeit solch eine Sensitivitätsanalyse durchzuführen ist Cross-Validation (CV).

¹⁹Hierbei wurden der Ausdruck des Geosplines (vgl. `Regression.R`) über ein Bild von Großbritannien gelegt und die Koordinaten angepasst. Da dies per Hand in Photoshop erfolgte, sind leichte räumliche Abweichungen möglich.

²⁰Vgl. „Tabelle 1 aus Papier“ und „Regr. aus Tabelle 1 ohne Geodaten“ in `regression.R`

Variante	additives Modell		simples Modell	
	adj. R2	AIC	adj. R2	AIC
1	0.684	131.1	0.655	131.4
2	0.695	130.6	0.647	133.2
3	0.676	132.9	0.652	132.5
5	0.813	88.7	0.613	110.8
6	0.810	89.5	0.603	112.6
7	0.865	74.3	0.602	112.7

Tabelle 7: Vergleich der Modelfits zwischen Regression mit und ohne Geodaten

CV ist ein einfacher Ansatz zur Schätzung der Genauigkeit und Vorhersagbarkeit von Regressionsmodellen. CV wird in vielen Bereichen des Maschinellen Lernens und der Regressionsmodellierung angewendet. Bei der hier verwendeten k-fold Cross Validaton handelt es sich um eine modifizierte Version mit teils selbst entwickelten Programmcode, dass auch mit GAMs umgehen kann. Die Vorgehensweise ist simpel: in mehreren Durchgängen werden die 41 Fälle in einen Trainings- und Testdatenteil aufgeteilt, mit denen jeweils die Regressionsmodellierung fünf mal wiederholt wird ($k=5$). Am Ende werden die Ergebnisse über alle Durchgänge gemittelt. Als Vergleichsstatistik wählte ich die die mittlere quadratische Abweichung (MSE). Die Ergebnisse für die beiden Regressionsmodelle sind in Tabelle 8 dargestellt. Man sieht, dass die MSE-Werte der Cross-Validierung deutlich höher sind, was grundsätzlich ein Hinweis darauf ist, dass das Regressionsmodell „overfitted“ ist. Solch ein Regressionsmodell entsteht normalerweise beim Trainieren großer Datensätze, bei denen die Regressionsgerade (meist mittels Gradient Descent) sich immer perfekter an die Trainingsdaten anpasst, hinsichtlich neuer Fälle (Testdaten) die zu erklärende Variable aber nicht prognostizieren kann. Bei der hier vorliegenden OLS-Regression mit nur 41 Fällen ist die Erklärung für das schlechte Abschneiden der Sensitivitätsanalyse nicht so klar zu benennen. Die empirischen Daten scheinen evtl. so gut zu passen, dass sie diese hohen MSE-Werte hervorbringen, eine kleine Variation der Daten aber schnell zu anderen Ergebnissen führt.

3.2.4 Logistische Regression

Da die Verteilungsfunktion der absolut Beschäftigten in der Textilindustrie stark rechts-schief ist, mit einigen wenigen Regionen mit sehr hoher und Regionen mit vergleichsweise geringer Textilindustrie - es fehlen quasi Regionen mit mittlerem Beschäftigungsniveau, stellt sich die Frage, ob eine binäre Kodierung der abhängigen Variablen nicht sinnvoller wäre, d.h. ob eine Klassifikation der Regionen in industrialisierte und (noch) nicht industrialisierte nicht bessere Ergebnisse liefern könnte.

Model		Papier	CV	CV ohne Geospline
tex1	Textiles_1831~Skills+Wage+Potential+SHP+s(X,Y)	0.870	2.139	1.885
tex2	Textiles_1831~Skills+Wage+Potential+Coal_Distance+SHP+s(X,Y)	0.800	2.071	1.975
tex3	Textiles_1831~Skills+Wage+Potential+Water_Flow+SHP+s(X,Y)	0.864	2.455	1.910
tex5	Textiles_1851~Skills+Wage+Potential+STF+s(X,Y)	0.173	1.220	0.957
tex6	Textiles_1851~Skills+Wage+Potential+Coal_Distance+STF+s(X,Y)	0.170	1.434	1.024
tex7	Textiles_1851~Skills+Wage+Potential+Water_Flow+STF+s(X,Y)	0.100	1.121	0.978

Tabelle 8: Vergleich der mittleren quadratischen Abweichung (MSE) zwischen Regressionsmodell aus Papier (Table 1) und k-fold Cross-Validation (CV)

	1831	1851
,„industrialisiert“	13	9
„(noch) nicht industrialisiert“	28	32
(n=41)		

Tabelle 9: Anteil industrialisierter Landkreise 1831 und 1851

Zur Überprüfung dieser Frage habe ich die Regressionsanalysen des Papiers als logistische Regressionsanalyse wiederholt. Hierzu wurden die Landkreise in zwei Gruppen geteilt: jene mit einer anteiligen Beschäftigung von mehr als 5% in der Textilindustrie und jene bis 5% (vgl. Tabelle 9). Trotz der Erwartung, dass die Regression einen guten „Fit“ hervorbringt, gab es keine signifikanten Regressionskoeffizienten. Verschiedene Variationen logistischer Regressionen wurden probiert. Das Ergebnis wirft Fragen auf: die Verteilung der Landkreise hinsichtlich Textilbeschäftigung weist klar auf eine Gruppenbildung hin. Das ein logistisches Regressionsmodell mit den gleichen Regressoren dennoch keinen Einfluss auf die binäre Variable Textilbeschäftigung zeigt, deutet darauf hin, dass die Regression der Autoren einen Großteil ihrer Erklärungskraft aus den leichten Variationen der Landkreise mit mittlerer Textilbeschäftigung hervorbringt. Die ausführliche Analyse ist auch hier wieder im R-Skript.²¹

²¹Vgl. Abschnitt „LOGISTISCHE REGRESSION“ in `Regression.R`

Gerechnet wurden nicht nur verschiedene Kombinationen von abhängigen Variablen, sondern auch traditionelle, ohne Korrektur des Geosplines, wie auch Varianten mit Korrektur.

4 Kritisch-hermeneutische Analyse

4.1 Aufbau des Papiers

Das Papier besteht aus 30 Seiten. Dazu kommen 5 Seiten Anhang, in denen auf 4 Seiten ausführlich auf die methodische Besonderheit der Regressionsanalysen eingegangen wird. Hier werden v.a. Simulationsergebnisse präsentiert, die übertrieben und etwas zu technisch wirken für den Rest des Papiers. Daneben wird im Anhang separat auf die Transportsituation in England eingegangen und die Herkunft der Daten beschrieben wird. Die Herkunft der Daten zu den wichtigen abhängigen Variablen (Textilbeschäftigung) wird zwar benannt, ist aber ohne Nachschauen im Working Paper (Kelly et al. 2015) kaum replizierbar. Das Verhältnis von Text zu Grafiken ist für eine empirische Arbeit ausgewogen.

4.2 Autoren

Die Autoren sind international angesehene Ökonomen. *Joel Mokyr*, ein in den Niederlanden geborener israelisch-amerikanischer Wirtschaftshistoriker an der Northwestern University (USA) kann wohl als führende Chorophäe hinsichtlich der Erforschung von Industrialisierung angesehen werden.

Morgan Kelly ist ein in Irland einem auch allgemeinem, nichtakademischen Publikum bekannter Wirtschaftshistoriker an der Universität Dublin mit Forschungsschwerpunkt auf irische Wirtschaftsgeschichte.

Cormac Ó Gráda ist emeritierter Professor für Wirtschaftsgeschichte ebenfalls an der Universität Dublin. Die Publikationsstatistiken der Autoren enthält Tabelle 10.²².

Darüber hinaus gibt es eine Danksagung für die vermutlich kritische Auseinandersetzung an u.a. *Peter Sola*²³, einem ebenfalls bekannten Wirtschaftshistoriker der WU Wien, dessen Ausführungen über Industrialisierung auch Inhalt einer hörenwerten Podcastfolge sind (vgl. Sator 2024) und in Kapitel 2 eingeflossen sind. Das Papier ist über einen Zeitraum von mindestens 7 Jahren entstanden.²⁴

Obgleich der Hauptautor *Joel Mokyr* hier eine strikt ökonomische Erklärung für Industrialisierung vertritt, bezeichnet Clark 2012 ihn als einen Vertreter des „idealistischen“ Ansatzes (als Gegensatz zu dem im Papier vertretenen materialistischen Ansatzes). Und in der Tat vertritt *Mokyr* die Meinung, dass Großbritannien den Rest Europas

²²Daten entnommen von semanticscholar.org am 10.08.2024.

²³Vgl. die Fußnoten auf S. 59f. in Kelly et al. 2023.

²⁴Es existiert ein Working Paper aus dem Jahr 2015 mit dem Titel „Roots of Industrial Revolution“ (Kelly et al. 2015).

Autoren	Publikationen	Zitationen	h-index
Joel Mokyr	453	13093	52
Cormac Ó Gráda	308	4406	36
Morgan Kelly	59	2922	27

Tabelle 10: Publikationsstatistiken

in die Industrielle Revolution geführt hatte, weil dort ein optimales „Klima“ aus Ideen, Kultur, Institutionen und Technologien bestand, die ein schnelles Wirtschaftswachstum ermöglichte. (Mokyr 2010)

4.3 Entstehungsdauer

Warum dauerte es so lange bis zur Publikation? Der erste Nachweis für ein Working Paper ist aus dem Jahre 2015 (Kelly et al. 2015). Das finale Papier (Kelly et al. 2023) wurde im Jahr 2022 akzeptiert und 2023 veröffentlicht. Natürlich kann nur gemutmaßt werden, warum zwischen Veröffentlichung des Working Papers und des Papiers mindestens sieben Jahre lagen. So könnte es der Tatsache geschuldet sein, dass die Autoren räumlich getrennt forschen und lehren (USA und Irland). Denkbar wäre aber auch, dass zwischen dem Working und finalen Papier noch viele Überarbeitungsschritte - evtl. während des Reviewprozesses oder sogar Ablehnung anderer Journals lagen. Oder die Autoren waren dermaßen überzeugt von ihrem empirischen Ergebnissen, dass sie ein möglichst großes Publikum durch die Wahl eines weniger fachspezifischen dafür allgemeineren Journals mit größerer Reichweite in der ökonomischen Community erreichen wollten.

4.4 Wahl des Journals

Das Papier wurde im JOURNAL OF POLITICAL ECONOMY publiziert. Dieses gehört zu den „Top Five“ der internationalen Journals im volkswirtschaftlichen Bereich. Zu fragen wäre, warum die Autoren nicht in einer stärker auf ihr Forschungsgebiet spezialisierten Zeitschrift publiziert haben. Zu denken wäre an das JOURNAL OF URBAN ECONOMICS oder das JOURNAL OF ECONOMIC HISTORY. Die starke Ausrichtung des Papiers auf die Regressionsmodellierung würde auch in das JOURNAL OF ECONOMETRICS passen. Durchaus möglich wäre, dass letzt genannte Journals - wenngleich (ebenso) renommiert - als „low-hanging fruits“ angesehen wurden und die vielversprechenden Ergebnisse der Regressionsmodellierung (ich erinnere hier noch mal an die ungewöhnlich hohen R-Werte) die Autoren motivierten, eine Veröffentlichung in einem Top-Five Journal an-

zustreben. Dies wiederum könnte das wahrgenommene Fehlen eines roten Fadens im Papier erklären.

Letztendlich wäre dies nur durch die Autoren selbst zu beantworten. Die Frage nach der Wahl des Journals bleibt somit ungeklärt, was aber bleibt ist die Feststellung, dass man ein allgemeines Publikationsmedium einem domänspezifischeren vorgezogen hat.

4.5 Wahl des Regressionsmodells

Begründung der Auswahl der Variablen Vergleichsweise viel Augenmerk im Papier wird der Begründung der Auswahl der erklärenden Variablen gewidmet. So wird eigens eine Regressionsanalyse vorgestellt, die den nicht-signifikanten Einfluss von Alphabetismus, Banken, Anwälten und Büchereien auf die Textilbeschäftigung darstellt (Tabelle 2 auf S. 78). Außerdem wird in einem zusätzlichen Kapitel im Anhang auf die Logistik eingegangen (S. 89). Eine ausführliche Erklärung dafür, warum die Textilbeschäftigung in den Jahren 1831 und 1851 ein geeigneter Indikator zur Messung von Industrialisierung ist, existiert hingegen nicht.

Begründung des Modells Ebenso auffallend ist der Stellenwert, der der Erklärung des Regressionsmodells gewidmet wird. Vermutet werden kann, dass die Verwendung dieser speziellen Regressionsform eine besondere Rolle im Begutachtungsprozesses des Papiers spielte. So wäre es denkbar, dass eine erste Version des Papiers noch multiple Regressionen enthielt - die Modelfitergebnisse waren ja durchaus beeindruckend²⁵ - zumal das Working Paper (Kelly et al. 2015) einige Hinweise auf andere Regressionsvarianten enthält. Möglich wäre, - aber das ist nur Spekulation - dass die Begutachter die Verwendung von Landkreisen als Untersuchungseinheit aufgrund der fehlenden Trennschärfe der Fälle kritisierten. So beginnen die Autoren den verhältnismäßig groß ausgefallenen Appendix über die Simulationsergebnisse ihrer semiparametrischen räumliche Regression mit: „*Regressions with spatial data frequently generate spuriously large t-statistics, ..*“ (S. 85). Es wäre durchaus denkbar, dass dieser Satz direkt aus dem ersten Gutachten stammt.

Eine alternative Erklärung wäre aber auch, dass die Autoren von Anfang das linear-additive Regressionsmodell verwendeten und die Gutachter eine bessere Darstellung dieses in Ökonomiekreisen vermutlich weniger bekannten statistischen Verfahrens forderten. Allerdings handelt es sich um im Data Analysis und Data Science Bereich durchaus weit verbreitetes Modell, das ebenso in der Wirtschaftsgeografie und Region-

²⁵Siehe die Ergebnisse der Regressionsanalyse ohne Geoplines im Abschnitt `Regressionen ohne Geodaten` in `regressions.R`

nalökonomie Verwendung findet. Darauf wiederum würde folgender Satz hindeuten: „*Despite their elegance, simplicity, and power, semiparametric regressions never took off in economics.*“ (S. 62f.)

4.6 Ungenaue Datenherkunft

Die wichtigsten beiden Variablen sind die Beschäftigung in der Textilindustrie für 1831 und 1851. Die Autoren schreiben „*Textile employment in 1831 is from Marshall (1833, 10–11).*“ (S. 89). Nun führte eine Nachverfolgung der Quelle²⁶ allerdings dazu, dass an beschriebener Stelle zunächst keine Textilbeschäftigungssdaten gefunden werden konnten. Das Buch besteht aus zwei Teilen, jeweils mit eigener Seitenzahlnumerierung. Somit gibt es zwei mal die Seiten 10-11. Ein Screenshot der beiden Seiten befindet sich im Anhang auf Seite 29. Der erste Teil enthält zwar Beschäftigungssdaten auf Landkreis-niveau für 1821 und 1831²⁷, jedoch keine Daten für die Textilindustrie oder ähnlicher Industrien.

Auch an anderen Stellen im Buch konnte kein Hinweis auf Textilbeschäftigungssda-taten gefunden werden. Einen Verweis darauf, wo die Textilbeschäftigungssdaten für 1851 kommen, gibt es im ganzen Papier nicht. Eine E-mail an einen der Co-Autoren ergab auch keine Antwort. Ein Hinweis des Betreuers während der Präsentation im Seminar nach einer Vorabversion (working paper) des Papiers zu suchen, ergab die schon mehr-fach erwähnte Quelle Kelly et al. 2015, in der zwar genauer beschrieben wird, wie die Daten für die Variable **Textile_1831** aus Marshall 1833, S. 10f. konstruiert wurde, mir aber dennoch nicht gelungen die Variable nachzubilden.

5 Fazit

5.1 Resümee

Stellt sich abschließend die Frage, wie das bisher geschriebene in eine finale Beurteilung des Papiers münden kann. Hierzu werde ich die Punkte die zugunsten der Vorgehenswei-se der Autoren und die, die dagegen sprechen nochmals gegenüberstellen (vgl. Tabelle 11). Es ergibt sich ein Gleichstand.

Auch wenn die Bewertung subjektiv ist und streng genommen kein abschließendes Urteil zulässt, würde ich doch den empirischen Ergebnissen und Schlussfolgerungen der

²⁶Marshall 1833

²⁷Der zweite Teil des Buches beginnt auf Seite 116 und beinhaltet detaillierte Steuer- und Zolleinn-hamen für verschiedene Waren.

Bewertungskriterium	positiv	negativ
Reputation der Autoren	+	
Reputation des Journals	+	
Roter Faden		-
Genauigkeit bei der Datenaufbereitung		-
Nachprüfbarkeit der Quellen		-
Datenbereitstellung	+	
Theoretisches Modell		
Wahl der Untersuchungseinheit		-
Komplexität des Modells		-
Empirisches Modell		
Teststärke	+	
Robustheit der Regressionskoeffizienten	+	
<i>Summe:</i>	5	5

Tabelle 11: Bewertung des Papiers

Autoren i.S. einer vorläufig gültigen Forschungshypothese, die im wissenschaftlichen Diskurs bestand halten muss, zustimmen. Es dürfte interessant sein, zu sehen, wie diese noch junge Veröffentlichung in Zukunft zitiert werden wird.

Somit kann man sagen: Industrialisierung ist also das Ergebnis niedriger Löhne und dem Vorhandensein von Fachkräften („mechanical skills“), die den technischen Fortschritt aufrecht erhalten. Noch abstrakter könnte man auch sagen: günstige Arbeitskräfte führen denmnach zu einem (komparativen) Kostenvorteil und spezialisierte Arbeitskräfte sorgen dafür, dass dieser Vorteil (durch technischen Fortschritt, bzw. Rationalisierung der Arbeitskräfte) bestehen bleibt.

Warum die industrielle Revolution im 18. Jahrhundert in England begann, kann aber auch das Papier nicht erklären. Immerhin liefert es aber einige interessante Erkenntnisse über die Entstehung der Textilindustrie während der beginnenden Industrialisierung in England. Die Verwendung von regionalen Daten auf Landkreisniveau ist pragmatisch aber auch mit Problemen behaftet.

Resümierend muss ich feststellen, dass es sich um ein interessantes Papier handelt, dass (neben den Forschungsergebnissen) auch wissenschaftssoziologisch sehr interessant ist. Es wäre interessant ein Experteninterview mit den Autoren zu führen, um den Entstehungsprozess besser zu verstehen.

5.2 Zusammenfassung

Mit dieser Seminararbeit versuchte ich, eine kritische Analysse des Papiers „*The Mechanics of the Industrial Revolution*“ vorzulegen. Dabei merkte ich, dass dies auf 20 Seiten kaum zu schaffen ist, die Ausführungen in den einzelnen Kapiteln sind dementsprechend knapp ausgefallen. Auch möchte ich betonen, dass, trotz aller Kritik an den Autoren, das Veröffentlichen der sauber aufbereiteten und mit ReadMe's versehenen Daten vorbildlich ist; eigentlich ein Standard, der für jede wissenschaftlichen auf Datenanalyse aufbauende Arbeit gelten sollte.

Ich war äußerst kritisch mit den Autoren, deswegen möchte ich in einem letzten Satz noch erwähnen, dass ich selber selbstverständlich auch nicht fehlerfrei bin: während der Überprüfung der Daten habe ich selber erfahren, wie viele Fehler (flüchtige, fahrlässige, aber auch gravierende) man machen kann, die dann durchaus zu falschen Schlussfolgerungen führen können. So musste ich selber mehrfach Vorabthesen, die ich zunächst als bestätigt sah, nachträglich verwerfen, da ich einen Fehler in meiner Datenanalyse entdeckt hatte. Und natürlich kann dies auch für einzelne Aussagen in dieser Seminararbeit nachträglich der Fall sein. Aus diesem Grund belasse ich das Github Repository zu dieser Seminararbeit auf unbestimmte Zeit öffentlich.

5.3 Nachtrag

Ursprünglich geplant und während des Seminars angekündigt, hatte ich ursprünglich vor, ein eigenes empirisches Modell zu entwickeln. Aus mehreren Gründen bin davon abgekommen:

- ★ eine kritische Anaylse und ein eigenes Modell wäre auf 20 Seiten nicht zu schaffen,
- ★ die vorgegebene Benchmark (hohe R^2 und niedrige AIC Werte) wären wohl kaum zu überbieten,
- ★ eine genauere räumliche Modellierung der Einflussfaktoren wäre statistisch zu anspruchsvoll gewesen.²⁸

Aus diesen Gründen habe ich mich darauf beschränkt, *lediglich* eine Idee, bzw. groben Entwurf für ein alternatives, kommunikationsbasiertes, systemtheoretisch-evolutionsbasiertes und dementsprechend stärker wirtschaftssziologisches Vorgehen vorzustellen (vgl. Kapitel 2.2).

²⁸Die meisten Daten sind ja Zensusdaten und auf Kommunalebene erheben worden. Dadurch könnte man kleinräumiger und somit genauer die Einflussfaktoren und deren Wechselwirkungen auf die Industrialisierung untersuchen.



Abbildung 4: The Industrial Revolution

Abschließen möchte ich die Seminararbeit mit Abb. 4 - einem Bild von *Cynthia Stokes-Brown*, dass ganz ohne Worte (und in Kontrast zu der Hauptaussage des in dieser Seminararbeit kritisch analysierten Papiers) die Bedeutung von Kohle in der Industrialisierung zeigt.²⁹

²⁹Entnommen aus: <https://images.app.goo.gl>

Literatur

Britain began industrializing in the 17th century - more than 100 years earlier than history books claim, 2024. URL <https://phy.org/news/2024-04-britain-began-industrializing-17th-century.html>.

Gregory Clark. The enlightened economy: An economic history of Britain 1700-1850: Review essay. *Journal of Economic Literature*, 50(1):85–95, 2012.

George Carrington Gray. *Gray's New Book of Roads*. Sherwood, Jones and Co., London, 1824.

Morgan Kelly, Joel Mokyr, and Cormac Ó Gráda. Roots of the industrial revolution. UCD Centre for Economic Research Working Paper Series, WP2015/24. University College Dublin. School of Economics, 10 2015. URL <https://researchrepository.ucd.ie/server/api/core/bitstreams/81f52f35-a676-4992-b166-e0d1eac46274/content>.

Morgan Kelly, Joel Mokyr, and Cormac Ó Gráda. The mechanics of the industrial revolution. *Journal of Political Economy*, 131(1):59–94, 2023. doi: 10.1086/720890. URL <https://doi.org/10.1086/720890>.

Khan-Academy. World history project - 1750 to the present. unit 3.1: Origins of the industrial revolution. <https://www.khanacademy.org/humanities/whp-1750>, o.A.

Malte Lierl. Systemtheoretischer evolutionsbegriff in geschichte und oekonomie, June 2004. Available at <https://core.ac.uk/download/pdf/236364518.pdf>.

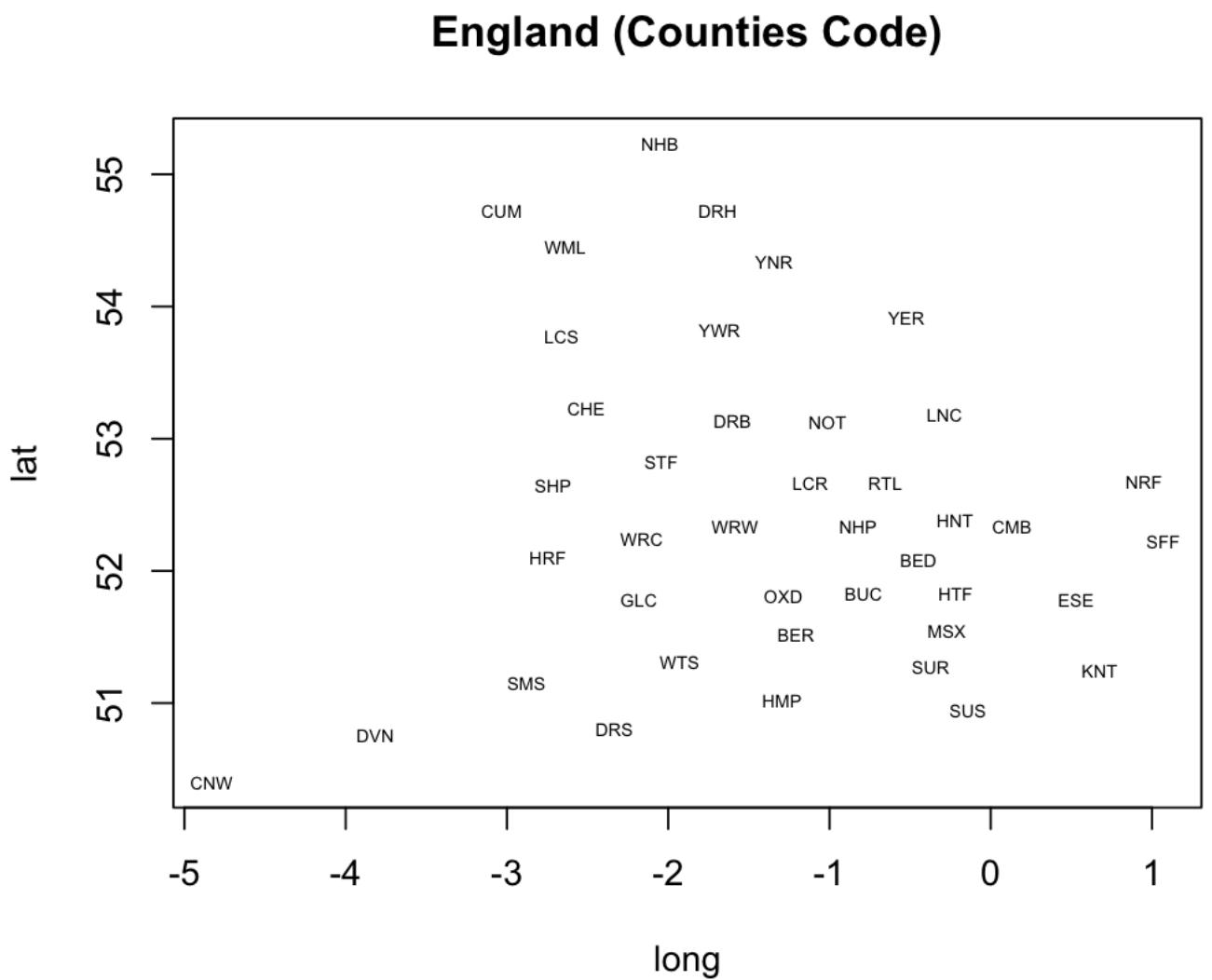
John Marshall. *A Digest of All the Accounts Relating to the Population, Productions, Revenues, [...] of the United Kingdom of Great Britain and Ireland: Diffused through [...] Journals, Reports, and Papers Presented to Parliament during the Last Thirty-Five Years*. J. Haddon, London, 1833.

Joel Mokyr. *The Enlightened economy an economic history of Britain 1700-1850*. Yale University Press, 2010. ISBN 9780300124552.

Andreas Sator. Erkläre mir die industrielle revolution, andreas resch. Podcast, Juni 2024. URL <https://xn--erklrmir-3za.at/2024/06/04/305-erklaer-mir-die-industrielle-revolution-andreas-resch/>.

A Anhang

A.1 Landkreise in England



Code	Name	Code	Name	Code	Name
BED	Bedfordshire	HNT	Huntingdonshire	NHP	Northamptonshire
BER	Berkshire	HRF	Herefordshire	SMS	Somerset
BUC	Buckinghamshire	HTF	Hertfordshire	STF	Staordshire
CHE	Cheshire	KNT	Kent	SUR	Surrey
CMB	Cambridgeshire	LNC	Leicestershire	SUS	Sussex
CNW	Cornwall	MSX	Middlesex	WML	Westmorland
CUM	Cumberland	NHB	Northumberland	WRC	Worcestershire
DRB	Derbyshire	NHP	Northamptonshire	WRW	Warwickshire
DRH	Durham	NOT	Nottinghamshire	WTS	Wiltshire
DRS	Dorset	NRF	Norfolk	YER	Yorkshire.East
DVN	Devon	OXD	Oxfordshire	YNR	Yorkshire.North
ESE	Essex	RTL	Rutland	YWR	Yorkshire.West
GLC	Gloucestershire	SFF	Suffolk		
HMP	Hampshire	SHP	Shropshire		

A.2 Variablenbeschreibung

Variablen	Beschreibung	Bedeutung	Min	Max	Mean	Outlier
Textile_1831	Anteil der männlichen Beschäftigten >20 Jahre in der Textilindustrie 1831	abhängige Variable	0.1	30.9	5.18	8
Textile_1851	Anteil der männlichen Beschäftigten 20-29 Jahre in der Textilindustrie 1851	abhängige Variable	0.1529	25.5959	4.2031	3
Potential	Marktpotential, d.h. Summe der aggregierten Einkommen	unabhängige Variable	12	661	126.3	1
population	Bevölkerung 1750???	unabhängige Variable	19601	1373400	320428	
wage	Löhne im Agrarbereich 1760. Logarithmierte Version der Variablen wage.1767	unabhängige Variable	72	117	85.24	3
skills	Fähigkeiten in der Metallverarbeitung	unabhängige Variable	23.31	117.26	54.12	2
Coal_Distance	<i>„... distance of the center of each county to the nearest county with a coal field. Counties with a coalfield were assigned a distance of 20 km.“</i> (S. 89)	unabhängige Variable	20	197	71.15	0
Water_Flow	<i>„(...) the area that drains into it (...). Each county is assigned a value equal to the 98th percentile of the flow across its squares.“</i> (S. 89)	unabhängige Variable	6	110	35.1	1
SHP	dummy Variable (true for Shropshire (SHP))	dummy Variable				

A.3 Fehlende Quelle für Textilindustriedaten

STATEMENT showing the Superficies in Statute Acres, and Rental, of each County of ENGLAND; the Number of Persons returned as dependant on Agriculture in 1801; the Number of Families so dependent at each of the three periods, 1811, 1821, and 1831; and the Number of Male Persons at the latter date, 20 Years of age and upwards effectively employed in Agriculture; shewing also the average of Superficies to each occupier, or Average extent of the Farms; of Labourers to each occupier; and the proportion of the Total to every 100 acres in each County.

The average extent of the Farms in Col. 12 has been deducted from Col. 2 on an average of 10 acres for each occupier not employing Labourers. In looking at the relative No. of Labourers in each County there are two things to be considered,—viz., the proportion of unproductive Land, and the greater or lesser extent of Tillage.

COUNTRIES arranged in order of their Extent of Superficies.	Superficies in Statute Acres, according to the Paragonometrical Survey.	Rental, Details of Par- ishes.	Persons returned as employ- ed in Agricul- ture,	Families returned as employed in Agriculture,	Male Agriculturalists, 20 years of age, and upwards, in 1831.				Total of the Superficies of England.	No. of Occupiers. Average extent of each Farm.	No. of Labourers. Average extent of each Occupier.	No. of Families to each Occupier.	No. of Servants employed.	No. of Servants employed.	No. of Servants employed.	
					1811.	1821.	1831.	Labourers.								
1 Lincoln	1,758,720	1,663,850	1,581,940	60,584	29,881	34,900	35,749	6,901	6,204	32,167	45,372	322	4,06	9,72	119	112
2 Devon	1,650,560	1,636,450	1,617,367	56,208	33,044	37,037	33,356	6,905	3,356	36,311	47,936	173	2,77	2,483	31	99
3 York, West Riding	1,651,000	1,629,890	1,655,603	55,615	30,868	31,613	31,108	7,006	10,636	24,062	41,284	214	3,46	2,63	142	172
4 Norfolk	1,338,880	1,302,300	981,812	31,464	21,406	21,701	37,610	5,229	2,718	7,16	24,413	241	3,51	53	104	104
5 York, North Riding	1,327,820	1,275,820	856,010	44,004	16,570	16,737	17,904	4,650	4,334	14,046	23,930	940	2,95	1,87	111	47
6 Northumberland	1,197,440	1,165,430	906,789	33,190	10,545	11,567	10,127	3,376	1,938	10,441	14,095	465	4,39	1,20	48	58
7 Lancashire	1,117,290	1,270,344	52,018	23,305	22,723	24,636	6,658	6,658	5,714	20,849	37,321	151	3,14	3,34	191	252
8 Somerset	1,050,080	1,028,090	1,315,100	61,434	27,472	31,418	30,422	6,032	3,731	28,107	37,970	164	4,65	3,69	71	116
9 Southampton	1,018,550	984,020	594,020	50,886	21,401	24,303	22,761	2,774	1,234	24,075	26,633	362	6,09	2,81	42	77
10 Kent	983,650	972,240	868,138	64,124	27,077	30,869	31,667	4,361	2,162	36,118	42,626	218	8,26	4,38	43	196
11 Essex	980,460	979,000	904,716	65,174	26,517	33,206	34,569	4,661	1,680	38,234	43,635	212	8,38	4,46	35	82
12 Suffolk	907,680	901,750	694,073	55,744	20,405	30,755	31,917	4,526	1,121	38,040	38,687	200	7,41	4,21	44	117
13 Cumbria	945,920	969,490	469,259	51,062	10,888	11,207	10,480	3,617	2,839	9,010	16,466	260	1,49	3,59	90	86
14 Sussex	936,320	907,920	549,960	38,925	19,778	21,929	22,450	3,160	1,330	36,125	30,615	263	0,95	1,37	31	87
15 Wilts	632,560	639,620	810,627	51,517	22,657	22,972	22,305	3,367	1,239	24,708	29,394	207	3,37	24	66	66
16 Salop	853,240	864,360	788,495	45,046	18,414	17,086	16,632	2,139	1,086	17,267	21,9	4,77	2,68	43	79	79
17 Cornwall	840,280	771,360	506,472	42,887	17,465	19,302	18,570	4,008	1,013	16,243	24,474	177	4,614	2,86	37	41
18 York, East Riding	{ 711,260	{ 52,440	{ 700,000	{ 31,530	{ 14,517	{ 15,400	{ 13,025	{ 3,321	{ 1,061	{ 12,727	{ 17,409	{ 190	{ 3,41	{ 2,47	{ 90	{ 43
19 City and County	{ 819,220	{ 52,440	{ 700,000	{ 31,530	{ 14,517	{ 15,400	{ 13,025	{ 3,321	{ 1,061	{ 12,727	{ 17,409	{ 190	{ 3,41	{ 2,47	{ 90	{ 43
20 Gloucester	734,720	736,290	765,035	43,350	18,361	18,295	18,156	3,731	3,049	16,927	26,448	270	6,65	2,75	25	48
21 Durham	679,040	679,530	507,063	31,200	18,217	18,288	9,427	8,408	2,229	15,644	24,242	185	4,44	3,29	60	97
22 Chester	673,200	649,050	676,064	30,923	16,396	16,120	16,397	4,374	4,050	16,094	23,827	130	3,45	3,02	110	119
23 Derby	650,040	663,160	621,688	51,743	14,263	14,582	13,324	4,257	10,563	16,170	166	3,19	2,73	66	73	73
24 Northampton	650,380	646,810	696,437	55,303	15,235	18,974	18,384	3,015	1,117	17,775	21,307	210	5,49	3,38	55	55
25 Dorset	632,200	627,220	489,025	23,294	12,982	14,821	14,601	2,243	1,077	14,056	17,286	270	6,65	2,75	25	48
26 Warwick	577,280	567,930	645,136	34,756	15,131	16,759	15,380	3,838	1,142	15,644	19,624	196	5,51	3,45	37	53
27 Hereford	550,400	543,800	453,907	31,500	12,199	12,556	12,500	2,605	1,077	12,213	19,361	249	3,04	2,01	38	55
28 Cambridge	549,120	530,820	483,265	28,054	12,831	15,636	16,003	2,421	1,206	15,988	19,385	216	6,48	3,61	14	60
29 Nottingham	535,680	525,800	524,392	23,293	13,029	14,634	14,634	2,414	1,179	17,799	21,365	181	4,46	3,20	30	65
30 Leicestershire	514,560	511,310	702,402	23,823	11,700	13,026	12,352	2,656	2,145	10,542	15,343	190	3,46	3,07	46	69
31 Westmoreland	488,320	485,960	291,556	12,141	4,613	5,006	4,455	1,435	1,635	3,474	6,594	326	3,42	1,33	50	32
32 Oxford	467,380	461,830	487,625	33,109	13,646	13,409	14,759	2,054	1,304	15,998	16,924	224	7,36	4,42	50	50
33 Berks	463,840	472,470	405,150	38,155	13,409	12,417	14,617	1,711	1,407	14,912	16,971	273	6,65	3,59	23	47
34 Surrey	485,120	474,380	389,901	2,746	12,417	14,944	14,617	1,873	1,417	16,761	16,761	249	8,94	4,08	12	39
35 Buckingham	473,600	463,820	498,677	26,083	13,938	16,640	16,383	2,152	1,453	16,748	19,348	213	7,73	4,17	43	62
36 Worcester	460,360	459,710	516,263	36,605	13,016	14,926	14,654	2,636	1,260	14,350	16,486	169	5,63	4,04	45	69
37 Hereford	387,920	400,370	12,350	20,611	11,998	13,405	13,268	1,518	1,399	14,700	16,617	261	4,61	4,15	12	38
38 Monmouth	318,720	324,310	203,676	12,377	5,316	6,020	5,614	1,648	1,143	4,846	7,636	169	2,93	2,35	16	28
39 Bedford	296,320	297,632	273,621	18,766	9,431	10,754	11,384	1,330	1,474	11,588	13,382	220	8,71	4,49	9	40
40 Huntingdon	266,000	241,650	209,076	9,556	5,361	6,455	6,221	1,657	1,337	5,987	7,221	270	6,95	2,98	10	62
41 Middlesex	180,400	179,590	349,142	13,417	9,088	9,383	9,082	1,030	490	11,376	12,946	166	10,63	7,19	14	62
42 Rutland	95,360	97,500	99,174	3,095	2,025	2,410	2,269	420	424	2,763	2,763	216	4,45	2,83	30	13
TOTALS.	32,322,400	31,770,615	27,890,350	1,282,227	607,353	773,732	761,348	11,160	94,833	744,407	980,560	295	5,25	3,08	2,34	3,134
Number of Columns.	1.	2.	3.	4.	5.	6.	7.	8.	9.	10.	11.	12.	13.	14.	15.	16.

A Comparative View of the Number of Families in each County in 1821 and 1831, returned under each of the three usual heads, viz.—1st, Employed in Agriculture; 2nd, in Trade, Manufactures or Handicraft; and 3rd, all others not included in either of the two preceding classes; and also a Statement of the Male Population in each County, 20 Years of Age and upwards in 1821, under each of six several heads of Occupation or Service; the Counties in the first instance arranged in order of their Total Number of Families, whereby the relative proportion of each class in each county is the more readily discernible, and in the other cases are arranged in reference to their Manufacturing Predominance, and the Manufacturing Population placed immediately in comparison with the Agricultural; another Statement exhibits the Male Population in column 11, under each of 140 heads of Occupation or Service in each County, so classified as to show their extent, relation, and importance to each other, and inconsiderability of the manufacturing operation in its external point of view, when compared with the interesting combination of County, 40 classified as to show their extent, relation, and importance to each other, and inconsiderability of the manufacturing operation in its external point of view.

Comparative View of the Population Returns of 1821 & 1831, and of the Occupations of the Population at the latter date.

COUNTIES, arranged in order of their Total Number of Families in 1831.	Proportion of FAMILIES, Returned as employed in										Male Persons in Years of Age and upwards in 1831, engaged in Occupations as under.											
	All others, not including in the two preceding					Trade, Manufactures, or Handicraft.					Agriculture.					COUNTRIES, arranged in reference to their Manufacturing Predominance.						
	1821.	1831.	1821.	1831.	1821.	1821.	1831.	1831.	1821.	1831.	1821.	1831.	1821.	1831.	1821.	1831.	1821.	1831.	1821.	1831.		
1 Middleses	1	314,939	9,682	161,936	173,822	91,122	130,335	49,457	22,549	163,920	73,755	97,517	37,321	97,517	1	Lancaster	97,517	37,321	97,517	37,321		
2 Lancaster	261,171	250,025	22,723	24,656	152,271	173,633	23,179	61,636	17,614	11,229	60,540	21,496	41,934	21,496	2	York, W. Ridg.	74,496	41,934	74,496	41,934		
3 York, West Riding	263,173	196,640	31,013	31,188	106,841	118,733	21,012	46,725	8,334	10,365	60,100	23,655	31,305	23,655	3	Chester	13,305	23,655	13,305	23,655		
4 Survey	161,466	106,077	14,944	14,647	46,811	49,616	21,051	44,814	14,235	10,765	44,139	24,878	14,250	16,856	4	Nottingham	14,250	16,856	14,250	16,856		
5 Devon	86,606	101,911	37,037	35,055	33,945	33,339	19,892	32,067	5,697	9,073	35,784	16,616	12,240	16,343	5	Derby	12,240	16,343	12,240	16,343		
6 Kent	83,399	97,142	30,989	31,907	29,180	29,419	24,899	36,056	5,503	14,570	34,257	15,245	8,633	16,170	6	Leicester	8,633	16,170	8,633	16,170		
7 Somerset	75,537	84,571	31,448	30,452	27,132	28,230	14,357	25,938	7,074	9,676	36,702	12,183	4,350	37,970	7	Somerset	4,350	37,970	4,350	37,970		
8 Norfolk	74,497	84,232	36,548	36,548	26,201	27,471	17,751	21,428	3,116	3,675	26,543	14,307	4,740	45,413	8	Norfolk	4,740	45,413	4,740	45,413		
9 Stafford	66,780	83,593	18,265	18,156	22,435	23,418	8,060	21,789	3,569	4,245	24,768	22,630	6,092	26,443	9	Gloucester	6,092	26,443	6,092	26,443		
10 Gloucester	72,156	75,907	23,170	21,185	35,390	37,179	13,079	29,082	6,047	6,568	22,716	19,716	3,497	23,334	10	Wills	3,497	23,334	3,497	23,334		
11 Durham	60,132	72,357	16,729	15,830	39,189	43,391	4,155	13,188	4,012	3,729	32,579	10,338	3,214	16,466	11	Cumberland	3,214	16,466	3,214	16,466		
12 Warwick	63,760	65,922	34,900	35,748	15,945	17,984	9,016	12,870	2,734	3,866	20,490	12,900	1,005	23,930	12	York, N. R.	1,005	23,930	1,005	23,930		
13 Lincoln	53,629	63,319	34,206	34,508	15,160	16,292	12,445	21,960	4,260	5,000	16,493	10,565	6,565	12,942	13	Stafford	20,755	24,242	20,755	24,242		
14 Essex	52,024	64,955	18,126	16,387	27,105	34,597	6,758	13,561	2,626	3,563	22,134	12,723	11,375	16,624	14	Warwick	11,375	16,624	11,375	16,624		
15 Southampton	57,042	64,652	18,310	20,903	22,761	24,303	13,829	20,903	3,784	3,947	23,164	13,348	8,024	10,456	15	Worcester	8,024	10,456	8,024	10,456		
16 Suffolk	55,064	61,582	30,765	31,491	17,418	18,116	6,351	11,929	2,228	3,490	18,107	10,336	8,203	7,636	16	Monmouth	8,203	7,636	8,203	7,636		
17 Cornwall	61,992	62,816	19,302	18,361	15,543	18,382	16,301	18,511	11,828	24,038	24,038	22,705	1,353	2,267	18	Durh m	1,353	2,267	1,353	2,267		
18 Durham	45,940	54,730	9,408	8,408	12,912	16,592	6,927	6,750	27,817	2,871	3,204	19,035	16,473	2,547	11,329	19	Northumberland	1,282	14,085	1,282	14,085	
19 Wilts	47,684	61,659	24,972	25,045	16,982	15,927	12,470	15,927	10,987	1,792	4,336	13,089	4,626	1,074	10,584	20	Westmoreland	1,074	10,584	1,074	10,584	
21 Sussex	43,565	51,216	22,450	21,920	17,920	18,489	15,463	17,777	6,737	13,121	1,948	19,208	8,161	2,164	21	Cornwall	10,064	12,916	10,064	12,916		
22 Derby	42,404	48,320	14,582	13,324	19,302	20,708	11,371	14,208	2,020	2,838	17,426	10,897	2,121	47,955	22	Middlesex	2,121	47,955	2,121	47,955		
23 Northumberland	43,193	49,364	11,567	10,187	20,565	20,446	14,246	10,396	23,991	2,567	3,228	17,149	13,939	2,476	42,636	24	Devon	2,476	42,636	2,476	42,636	
24 Worcester	39,006	46,512	14,926	14,654	15,656	15,630	5,514	11,828	2,085	2,835	13,683	6,544	3,53	5,476	25	Kent	1,074	5,476	1,074	5,476		
25 Nottingham	35,603	43,603	13,408	13,684	14,632	25,676	9,183	14,632	9,183	9,093	2,320	14,663	6,628	2,313	167	Lincoln	167	45,279	167	45,279		
26 Leicestershire	38,803	42,142	13,027	12,352	20,287	22,984	3,481	6,860	1,422	2,509	13,772	3,701	2,717	3,701	27	Esser	2,717	3,701	2,717	3,701		
27 York, North Ridg	31,721	31,731	17,964	17,964	11,700	11,700	11,700	11,700	1,493	1,370	12,470	4,391	2,619	2,619	28	Southampton	2,619	2,619	2,619	2,619		
28 Northampton	35,552	39,163	10,974	11,384	12,866	12,895	4,883	7,434	2,084	2,874	13,841	8,766	3,087	3,087	29	Suffolk	3,087	3,087	3,087	3,087		
29 York, East Ridg	40,499	36,960	14,485	14,485	16,210	17,320	10,825	10,825	1,382	2,388	12,927	8,603	2,916	2,916	30	Sussex	2,916	2,916	2,916	2,916		
31 City & Anstray	31,804	34,820	11,297	10,630	13,146	12,026	7,961	12,164	1,202	1,202	9,130	6,136	6,077	6,077	31	Northampton	582	16,907	582	16,907		
32 Dorset	30,312	33,614	14,821	14,821	14,601	10,311	10,106	4,880	8,907	1,624	1,624	10,568	4,461	1,474	1,474	32	York, E. Ridg	1,474	17,600	1,474	17,600	
33 Bucks	26,867	31,849	16,640	16,393	16,393	8,395	3,395	6,561	1,176	1,820	8,694	3,213	3,213	3,213	33	Dorset	3,213	1,777	3,213	1,777		
34 Oxford	26,641	31,770	15,965	15,304	9,454	9,454	7,012	2,254	2,254	2,254	11,110	3,049	3,049	3,049	34	Bucks	3,049	3,049	3,049	3,049		
35 Berks	27,700	31,631	14,789	14,047	8,773	9,384	4,168	1,447	1,447	1,447	10,768	3,708	3,708	3,708	35	Gloucest.	3,708	3,708	3,708	3,708		
36 Hertford	26,170	30,210	13,465	13,268	7,635	8,532	4,750	7,130	1,629	1,629	9,426	4,101	711	18,510	36	Orford	711	18,510	711	18,510		
37 Cambridge	25,603	26,250	20,116	19,088	6,984	6,984	6,213	6,404	1,704	1,704	6,702	3,007	3,007	3,007	37	Beds	521	16,971	521	16,971		
38 Hereford	21,917	23,565	13,568	12,633	6,105	6,105	5,726	4,572	911	1,521	7,576	2,410	3,677	2,410	38	Hertford	2,410	16,617	2,410	16,617		
39 Bedford	17,373	20,016	10,754	11,364	5,137	5,137	3,515	5,577	1,037	1,037	5,592	1,474	1,474	1,474	39	Hereford	1,474	16,397	1,474	16,397		
40 Monmouth	14,132	19,911	6,029	5,614	6,147	8,626	1,955	5,671	805	1,011	6,649	7,173	7,173	7,173	40	Bedford	7,173	17,266	7,173	17,266		
41 Huntington	10,897	11,976	6,455	6,231	2,937	2,940	1,025	2,107	401	639	3,443	971	971	971	41	Bedford	971	17,392	971	17,392		
42 Westmoreland	10,438	10,984	5,036	4,454	3,301	4,454	1,022	1,022	1,022	1,022	6,621	1,106	1,106	1,106	42	Huntingdon	1,106	17,221	1,106	17,221		
43 Rutland	3,936	4,191	2,410	2,250	1,084	1,084	1,102	492	790	150	224	1,373	209	209	43	Rathlend	12	2,763	12	2,763		
TOTALS.	2,346,717	2,745,236	2,753,332	2,761,348	1,135,895	1,187,925	484,650	96,076	179,985	183,595	964,577	500,950	TOTALS.	500,950	500,950	500,950			314,105	900,540		

Eidesstattliche Erklärung

Ich erkläre, dass ich die Seminar-/Bachelor-/Masterarbeit selbstständig und ohne unzulässige Inanspruchnahme Dritter verfasst habe. Ich habe dabei nur die angegebenen Quellen und Hilfsmittel verwendet und die aus diesen wörtlich, inhaltlich oder sinngemäß entnommenen Stellen als solche den wissenschaftlichen Anforderungen entsprechend kenntlich gemacht. Die Versicherung selbstständiger Arbeit gilt auch für Zeichnungen, Skizzen oder graphische Darstellungen. Die Arbeit wurde bisher in gleicher oder ähnlicher Form weder derselben noch einer anderen Prüfungsbehörde vorgelegt und auch noch nicht veröffentlicht. Mit der Abgabe der elektronischen Fassung der endgültigen Version der Arbeit nehme ich zur Kenntnis, dass diese mit Hilfe eines Plagiatserkennungsdienstes auf enthaltene Plagiate überprüft und ausschließlich für Prüfungszwecke gespeichert wird.

Ort, Datum

Unterschrift