# grajzl_2019_toward_understanding_17th_century_english_culture_a_structural_topic_model_of_francis_bacons_ideas

## Year

2019

## Author(s)

Peter Grajzl and Peter Murrell

## Title

Toward understanding 17th century English culture: A structural topic model of Francis Bacon's ideas

## Venue

Journal of Comparative Economics

---

## Topic labeling

Manual

## Focus

Secondary

## Type of contribution

Established approach

## Underlying technique

Manual labeling assisted by associated documents

## Topic labeling parameters

Examined associated documents: 20

# Label generation

For each estimated topic, we examined closely the twenty documents that featured a given topic most prominently.

**Table 3**

Topics and top words for the estimated 16-topic STM.

**1. Epistemology**
*Highest prob*: natur, man, will, thing, philosophi, may, histori, knowledg, scienc, part, use, one, experi, can, mind, great, upon, shall, make, therefor, yet, first, matter, particular, understand, caus, much, invent, observ, true
*FREX*: philosophi, method, defici, logic, system, scienc, axiom, histori, metaphys, idol, discoveri, everyth, aristotl, invent, inquiri, theori, abstract, mathemat, rhetor, notion, mankind, deduc, knowledg, induct, fallaci, theolog, inventor, principl, poesi, contempl

**2. Probing for Facts**
*Highest prob*: say, will, man, one, shall, upon, come, great, answer, lord, make, take, give, king, may, think, day, tell, know, ask, hous, can, sir, see, time, like, use, much, mani, friend
*FREX*: overburi, impoison, somerset, ask, weston, diogen, galleri, bensalem, tell, poison, answer, pompey, nichola, vespasian, dinner, aristippus, gentleman, injunct, sir, front, chamber, yes, forgiv, tower, philip, friend, chair, ladi, feast, boat

**3. Land Law**
*Highest prob*: use, law, shall, statut, make, feoff, upon, case, will, word, take, land, yet, therefor, may, seiz, estat, heir, first, can, act, reason, say, good, one, time, life, give, right, grant
*FREX*: feoff, cesti, feme, lesse, remaind, disseise, statut, aver, seiz, clausula, rend, duress, usuri, claus, leas, purchas, advowson, remitt, commonlaw, fee, proviso, tail, parol, demis, atturn, void, dissent, deed, bargain, enrol

**4. King, Court & Law**
*Highest prob*: majesti, will, shall, may, good, lord, upon, time, lordship, king, make, think, self, man, can, know, great, say, take, matter, letter, give, much, well, thing, god, now, mean, one, first
*FREX*: lordship, plantat, majesti, humbl, undertak, letter, squir, essex, solicitor, attorney, gracious, lord, project, pleas, queen, secretari, coke, chancellor, servic, busi, con, honest, advertis, pray, manuel, antonio, wish, care, protest, beseech

**5. Law & Nation**
*Highest prob*: law, king, shall, may, england, will, natur, majesti, one, subject, case, say, make, time, first, kingdom, man, therefor, can, scotland, person, upon, court, question, sever, point, whether, parliament, take, yet
*FREX*: scotland, england, duel, allegi, born, union, parliament, kingdom, britain, challeng, alien, conquest, gascoign, subject, question, vote, nation, offenc, royal, law, normandi, post, preced, crown, submiss, prerog, duchi, infer, court, style

**6. Religious Law**
*Highest prob*: god, man, time, learn, law, may, great, make, church, will, shall, upon, good, work, one, say, yet, thing, see, majesti, well, give, first, natur, matter, part, can, use, mani, person
*FREX*: learn, church, christ, preach, preacher, atheism, benefic, scriptur, dedic, controversi, baptism, liturgi, hospit, endow, alexand, god, theme, holi, reveal, prayer, ecclesiast, bishop, institut, ministri, minist, callisthen, censur, pastor, polici, compil

**7. Legal Theory**
*Highest prob*: land, man, shall, king, good, feloni, may, make, upon, court, treason, call, take, justic, heir, parti, give, one, can, peac, will, law, offic, constabl, everi, yet, life, lord, grant, person
*FREX*: feloni, constabl, sheriff, counti, debt, executor, forfeit, escheat, juri, tenur, legaci, attaint, treason, petti, assiz, shire, felon, leet, attaind, gaol, owner, chattel, chap, indict, viii, clerk, writ, clergi, outlawri, praemunir

**8. Religion & Diplomacy**
*Highest prob*: great, upon, war, will, man, state, make, spain, time, say, shall, yet, may, peopl, part, one, king, take, come, now, england, mani, law, much, like, natur, nation, true, never, princ
*FREX*: spain, invas, spaniard, speaker, libel, cathol, war, spanish, enterpris, pollio, christendom, turk, sundri, heresi, leagu, palatin, navi, indi, defens, portug, papist, religion, nation, invad, fleet, confeder, armi, worship, lowcountri, nobil

**9. Political Strategy**
*Highest prob*: man, good, will, make, great, say, thing, upon, one, may, natur, can, time, much, shall, mind, like, yet, see, well, mani, come, part, virtu, take, therefor, use, person, first, fortun
*FREX*: envi, anger, dissimul, malum, felic, bewar, faction, evil, bold, tacitus, reprehens, cun, solomon, lover, secreci, fortun, gradus, discontent, reprehend, simul, virtu, precept, tiberius, friendship, demosthen, quoth, cicero, busi, machiavelli, eam

**10. Dynastic Politics**
*Highest prob*: king, upon, make, great, time, will, shall, come, man, may, part, take, person, one, peopl, also, yet, unto, war, well, princ, england, think, good, lord, much, can, duke, give, peac
*FREX*: perkin, maximilian, duke, earl, plantagenet, ferdinando, york, french, flander, margaret, ambassador, treati, richard, castill, charl, britain, ladi, rebel, ambassag, king, fillip, henri, lovel, london, john, duchess, bruge, stanley, tower, castl

**11. Classical Thought**
*Highest prob*: thing, man, may, natur, also, shall, say, seem, god, see, will, one, yet, unto, make, matter, can, time, great, like, take, come, mani, first, kind, well, fabl, much, jupit, upon
*FREX*: prometheus, jupit, proserpina, page, orpheus, miss, pan, cupid, siren, typhon, pentheus, sphinx, parabl, albeit, perseus, moreov, fabl, nemesi, giant, palla, bacchus, cere, muse, allegori, atalanta, monster, hercul, proteus, hell, icarus

**12. Physics, Energy**
*Highest prob*: motion, heat, bodi, natur, instanc, water, will, air, also, must, may, fire, place, can, part, let, one, cold, first, like, flame, yet, substanc, great, spirit, differ, observ, thing, power, appear
*FREX*: magnet, instanc, expans, similar, class, howev, anim, rapid, latent, liquid, heat, predomin, ignit, exist, ebb, homogen, heterogen, migrat, perpendicular, negat, expand, conspicu, investig, ray, excit, concret, motion, sphere, flame, exclus

**13. Physics, Air & Sound**
*Highest prob*: wind, sound, air, will, make, one, may, blow, thing, great, motion, water, part, bodi, upon, see, like, likewis, also, much, sail, caus, place, sea, man, come, shall, yet, south, two
*FREX*: wind, string, sound, brass, sail, rain, tone, blow, south, echo, percuss, bell, mast, east, pipe, nois, north, presag, articul, concav, audibl, dram, lute, west, cloud, nurseri, gale, loud, engend, tin

**14. Botany**
*Highest prob*: tree, will, upon, plant, fruit, make, may, earth, water, also, put, come, ground, grow, root, caus, forth, like, great, herb, seed, flower, set, leav, see, kind, bear, one, therefor, much
*FREX*: sap, plum, cherri, tree, moss, oak, pear, herb, bough, holli, stalk, peach, dung, radish, fig, graft, blossom, cucumb, plant, lettuc, rosemari, mushroom, wheat, vine, escul, colewort, compost, mistleto, turnip, fruit

## Motivation

The assignment of names to the estimated topics is a crucial step in the analysis because our ultimate conclusions about the features and the cultural origins of Bacon's thought are based on which particular topics are reflected in Bacon's corpus, which topics are correlated across documents, which topics entail overlapping vocabulary, and how Bacon's usage of different topics varies with both intended audience and document form.

Readers will note below that we are able to easily identify the underlying ideas for all 16 topics and assign non-controversial names for each. This is evidence for the internal validity of our results.

## Topic modeling

STM

## Topic modeling parameters

Nr of topics: 5 to 25

## Nr. of topics

16

## Label

Single or multi-word manually assigned labels

## Label selection

\

# Label quality evaluation

\

## Assessors

\

---

## Domain

Paper: Law, History
Dataset: Law

## Problem statement

We use machine-learning methods to study the features and origins of the ideas of Francis Bacon, a key figure who provided the intellectual roots of a cultural paradigm that spurred modern economic development.

Bacon's works are the data in an estimation of a structural topic model, a recently developed methodology for analysis of text corpora.

The estimates uncover sixteen topics prominent in Bacon's opus. Two are key elements of the ideas usually associated with Bacon—inductive epistemology and fact-seeking.

We provide the first quantitative evidence that the genesis of Bacon's epistemology lies in his experience in the common-law.

Combining our findings with accepted arguments in the existing literature, we suggest that the effects of common-law culture can help explain the coincidence of early political and economic development in England.

## Corpus

Origin: Project Gutenberg, Hathi Trust, Internet Archive, The Text Creation Partnership for Early English Books Online (2014).
Nr. of documents: 282
Details:

- all Bacon's works that had been digitized accurately enough that most words could be machine read.

**Table. 1**
Works included in the corpus.

| Name or type of work | No. of text documents (chunks) |
|---|---|
| *New Atlantis* | 5 |
| *Novum Organum* | 17 |
| *The Advancement of Learning* | 12 |
| *The Use of the Law* | 5 |
| *A Collection of Apothegms, New and Old* | 6 |
| *History Natural and Experimental of Life and Death* | 9 |
| *Sylva Sylvarum* | 10 |
| *The History of The Reign of King Henry the Seventh* | 6 |
| *The Natural and Experimental History of Winds* | 7 |
| *The Wisdom of the Ancients* | 5 |
| *The Elements of the Common Laws of England* | 28 |
| Other (various essays, letters, speeches, case reports) | 172 |
| Total | 282 |

*Notes*: Some names have been edited for rendering in modern English.

**Table 2**
Frequency of documents by audience and form of communication.

| Form of communication Audience | Apothegm | Case report | Essay | Letter | Speech | Tome | Total |
|---|---|---|---|---|---|---|---|
| Historians | 0 | 0 | 2 | 0 | 0 | 12 | 14 |
| Lawyers | 25 | 1 | 12 | 22 | 10 | 4 | 74 |
| Methodologists | 1 | 0 | 3 | 1 | 0 | 34 | 39 |
| Philosophers | 6 | 0 | 42 | 2 | 0 | 0 | 50 |
| Politicians | 2 | 1 | 16 | 36 | 3 | 0 | 58 |
| Scientists | 0 | 0 | 20 | 0 | 1 | 26 | 47 |
| Total | 34 | 2 | 95 | 61 | 14 | 76 | 282 |

## Document

Document (chunk) part of one of Bacon's literary work + two metavariables, namely:
- Bacon's intended audience—historians, methodologists, lawyers, politicians, scientists, or philosophers
- Form of his finished work—case reports, apothegms, letters, essays, book-length tomes, or speeches.

Topical prevalence (or importance) is modelled as a function of a linear combination of the dummy variables reflecting the two metadata covariates. Thus, the $K$ topics that we identify with Bacon are the same across all documents but the prevalence of each topic (element of d) varies with document type.

## Pre-processing

- The chaotic orthography of late 16th and early 17th century English was converted into standard modern orthography using a program that absorbed the database available with MorphAdorner.
- word inflections were then removed by converting all words to their lexical roots (stemming)

- latin words were stemmed and translated into English
- all words were converted to lower case
- Standard English stop words (natural language words which carry very little meaning, such as 'and', 'the', 'a', 'an'), numbers, and punctuation were removed.

---

```
@article{grajzl_2019_toward_understanding_17th_century_english_culture_a_struct
ural_topic_model_of_francis_bacons_ideas,
    abstract = {We use machine-learning methods to study the features and origins
of the ideas of Francis Bacon, a key figure who provided the intellectual roots
of a cultural paradigm that spurred modern economic development. Bacon's works
are the data in an estimation of a structural topic model, a recently developed
methodology for analysis of text corpora. The estimates uncover sixteen topics
prominent in Bacon's opus. Two are key elements of the ideas usually associated
with Bacon---inductive epistemology and fact-seeking. The utilitarian promise
of science and the centralized organization of the scientific quest, embraced
by Bacon's followers, were not emphasized by him. Using strategic
communication, Bacon facilitated reception of his scientific methodology,
targeted influential groups, and finessed powerful opponents. We provide the
first quantitative evidence that the genesis of Bacon's epistemology lies in
his experience in the common-law. Combining our findings with accepted
arguments in the existing literature, we suggest that the effects of common-law
culture can help explain the coincidence of early political and economic
development in England.},
    author = {Peter Grajzl and Peter Murrell},
    date-added = {2023-04-01 18:58:24 +0200},
    date-modified = {2023-04-01 18:58:24 +0200},
    doi = {https://doi.org/10.1016/j.jce.2018.10.004},
    issn = {0147-5967},
    journal = {Journal of Comparative Economics},
    keywords = {Francis Bacon, Culture, England, Law, Knowledge, Scientific
method},
    number = {1},
    pages = {111-135},
    title = {Toward understanding 17th century English culture: A structural
topic model of Francis Bacon's ideas},
    url = {https://www.sciencedirect.com/science/article/pii/S0147596718304426},
```

```
    volume = {47},
    year = {2019}}
```

#Thesis/Papers/BS