The Levellers in Context
Equals, Peers and Free-born Englishmen[*]

Melissa Schwartzberg[†]   Arthur Spirling[‡]

10407 Words

Abstract

We consider the evolution of Leveller thought relative to their contemporaries during the English Civil War(s). We compile a new data set of hundreds of seventeenth century pamphlets to chart themes related to rights, elections and suffrage for a variety of actors for the period 1638–1666. We combine this with novel word embedding techniques trained on millions of *Early Modern English* documents to make statements about how actors of the time understood key issues of interest to modern democratic theory. Our quantitative and qualitative results show that *inter alia* (1) John Lilburne created the key rhetoric of "freeborn Englishman" but that its evolution is more complex than assumed; (2) the Leveller interest in "equality" is distinctively legal rather than "social" in nature; (3) this legal conception also extends to Leveller concerns with elections, but on this topic they become more aligned with their contemporaries over time than in the case of equality.

1 Introduction

The Levellers have long served as a mirror for the ideological commitments of their interpreters. A political group active during the English Civil War(s) of the mid-seventeenth century, the Levellers are variously characterized by historians as "proto-democrats," advocates of an expanded franchise, legal rights, and religious toleration; as derivative theorists of the "Norman Yoke" of the common law; as bourgeois individualists or radicals manqué. Despite, or perhaps because, of these potentially contradictory readings, they remain seductive as a subject of study and reverence. Indeed, in the modern UK, partisan forces from both the communitarian left (Tony Benn, Jeremy Corbyn) and the libertarian right (Daniel Hannan, Douglas Carswell) cite them as inspirations. Today, the *Putney Debates* of 1647—for which three days of transcripts were rediscovered in 1890—are a staple of courses in modern political thought and democratic theory. A fragment from these meetings by Colonel Thomas Rainsborough is so familiar as to constitute a cliché of Leveller ideology: "I think that the poorest he that is in England hath a life to live, as the greatest he; and therefore truly, sir, I think it's clear that every man who is to live under a government ought first by his own consent to put himself under the government; and I do think that the poorest man in England is not at all bound in a strict sense to that government he has not had a voice to put himself under."

While the Levellers have been of continual interest to political theorists (e.g. Anderson, 2017; Bejan, 2022; Macpherson, 1962; Waldron, 2016), studying them systematically is difficult. A major issue is data availability and awareness: the Levellers wrote a large number of pamphlets and other material, only some of which are widely anthologized and have received much scholarly attention. We would expect a broader corpus of Leveller works to be more heterogenous—and potentially representative—than a selective reading of their "major" works. Beyond this, without documents from other authors of the period, it is hard to identify those arguments that are distinctive to the Levellers, as opposed to other factions of the Civil War era. Related, exactly who counts as a Leveller at a given time may be debatable since they "did not make up a structured and organized political movement fashioned by authors understood to be party leaders." (De Krey, 2017,

2) Third, within the subset of recognized Levellers, taking any argument or figure to be representative of their ideology as a whole is perilous: the writings of Lilburne pose particular challenges (Foxley, 2004; Foxley et al., 2016). That is, it is difficult to assess whether claims advanced by one or two major actors reflected, or produced, agreement among the Levellers as a whole, or were idiosyncratic to these individuals.

These are political theory problems for which quantitative text-as-data approaches are tailor made (see, e.g., Blaydes, Grimmer and McQueen, 2018). In keeping with this, one of contributions below is to use a new, relatively large and comprehensive corpus of pamphlets from the seventeenth century that combines "core" Leveller works with those of contemporaries. This data, which comprises some 260 contemporary documents of various lengths,[1] is derived from the *Leveller Tracts* project (Hart and Kenyon, 2018). It covers the years 1638–1666, and includes the social and political writings of over 60 different authors from that period. In that sense, it is a broader, deeper collection than previously employed. This allows greater purchase on the issues of comparison and purported distinctiveness we mentioned above. Of course, such a dataset is only useful to the extent that it can be analyzed appropriately. Our second contribution then is to provide and use new tools for this purpose. In particular, alongside some basic keyword matching, we will employ *word embedding* approaches. These techniques are proving popular in political science when researchers want to capture the "meaning" of terms, across groups or over time (Rheault and Cochrane, 2020; Rodman, 2020). They do this by implicitly incorporating information about the contexts—other words—of the terms of interest. We will use a specific variant of these models that allows for statistical testing inferences, even when we have relatively little data (i.e. when the terms are quite rare) (Rodriguez, Spirling and Stewart, 2021*a*). This arrangement nonetheless requires a background corpus written in a style and period approximating that of the pamphlets. Here we will use millions of words from official reports of parliament in the seventeenth century (courtesy of Rodon and Paskhalis, 2021). This leads us to build up embeddings for (political) *Early Modern English*, that we believe have not previously existed and that we anticipate will be *per se* useful to

---

[1] Technically, there are 262 documents, but two are repeats.

other scholars.

The combination of novel data and methods allows us to shed new light on some classic questions about the Levellers that have of been of intense attention in recent times. Broadly, these concern the supposed *social egalitarianism* of Leveller doctrine and the *relative importance of legal rights* in Leveller thought. On the former, we investigate the empirical evidence for political philosophy accounts of "relational" equality that pertain to "status". Specifically, we clarify the development of the use of "free-born" in Leveller writings, with an eye to the plural noun it modifies, from "subjects" to "Englishman." While we agree with previous literature that this appears to be an innovation of John Lilburne's, we refine this understanding to show that the shift was not once-and-for-all, and that it had resonance—and was picked up by authors—outside the core Leveller group. We also distinguish Leveller claims about social equality from those defended by others in the period. To that end, we assess the importance (frequency and spread) of other phrases taken to be significant by contemporary scholars, "equals and alike" and "equals or peers," both within and outside of Leveller texts. Here, the Leveller commitment to social equality reveals itself to be muted and statistically significantly different in tone to others, particularly the Diggers.

Given these findings, we ask what type of equality, if any, the Levellers *were* invested in. The answer is legal freedoms, broadly construed. In particular, the distinctiveness of Leveller thought as a whole rests on their commitment to an account of legal rights, from the power of Parliament to engage in deliberate legal change to rights to a fair trial to jury service. As we show, Lilburne is vital and influential to this position, though it is picked up by wider body of Levellers, and endures over time. By contrast, non-Levellers of the period tend not to rely on legal argumentation. Our final set of findings concern electoral institutions, and the difference between Levellers and their contemporaries's discussions of the same. In keeping with our findings for equality, we show that the Levellers understood the machinery and practice of elections in terms of legal freedom, and that this is distinctive from non-Levellers who do not use this lens. Interestingly though, the time series profiles for equality and elections are not the same: we provide evidence that while the Leveller understanding of the former never becomes widely adopted by others of the period,

there is considerably more convergence in discussions of democratic institutions—especially at the time of the publication of the final proto-constitution of the *Agreement of the People* series (May 1649). Put otherwise, we show that what makes the Levellers 'special' is not an early or unusual commitment to democracy, it is an early or unusual commitment to a particular idea *underlying* democracy.

In the next section we briefly situate the Levellers as an entity in history, before discussing 'live' debates on their beliefs and doctrine. This sets up a series of research questions for which our new data—described in Section 3—is arguably ideal. We introduce our embedding methods in Section 4 and provide some validation for the same, focussing on the challenges of working with *Early Modern English*. Our empirical results are presented in 5, and we conclude in Section 6.

2   The Levellers, Egalitarianism and Legal Liberties

Although Rainsborough's "poorest he" speech is perhaps the most-recited summary of Leveller thought, and the *Putney Debates* mark the beginning of the "zenith" of Leveller power (Frank, 1955, 135), the group's practical influence on England's constitutional settlement after the first Civil War—especially as regards suffrage reform—was minimal. Indeed, the Levellers were essentially an extinguished force by 1650. But they remain of both political and historical interest for at least two reasons. First, the Levellers arguably represent a genuinely new form of radicalism (Foxley, 2004). Their central ideas—of freedom of conscience, constitutional reform and parliamentary sovereignty—broke new ground in democratic thought. Second, the Levellers communicated widely using the relatively new technology of printing. That is, we have an unusually dense record of information about their doctrine, discussions, and how they evolved.

In recent years, the democratic character of Leveller thought has focused on its egalitarianism. Scholars have emphasized its anti-hierarchical strains. In her first Tanner Lecture, Elizabeth Anderson turns to the Levellers to vindicate a conception of social or relational equality (Anderson, 2017). For Anderson, egalitarianism, including the egalitarianism of the Levellers, is a means for

5

"dismantling or taming social hierarchy." (Anderson, 2017, 8) Specifically, she argues that the Levellers challenged the authority of priests and the Church of England through defending toleration; the domination of men over women through expanded franchise; and the power of guild monopolies through advocacy of private property and free trade. Anderson highlights the persuasive (and on her account, feminist) force of Leveller John Lilburne's characterization of Adam and Eve and their progeny as "by nature all alike and equal in power, dignity, authority, and majesty, none of them having by nature any authority, dominion, or magisterial power one over or above another."

In a recent article, political theorist Teresa Bejan modifies Anderson's argument slightly, arguing that the concept of natural "parity"— peerage among a set of free-born English men—better characterizes the Leveller commitment to equality. In Bejan's account, the notion that human beings are equal by nature—that they were, in her language, "indifferent" in God's eyes—had been a religious and juridical commonplace for centuries. The distinctive contribution of the Levellers then was to create a rank of men who possessed equal high standing.[2] Bejan holds that whereas the Levellers rejected natural hierarchies among human beings, they did also accept natural *differences* among them (e.g., between the sexes) and allocated rights unequally according to these disparities (Bejan, 2022, 605). For Bejan, even the concept of a "free-borne Englishman" was one of peerage, adopted by Lilburne and his associates to expand a set of legal rights previously held only by high-status "peers" to all subjects of England. Supporting Anderson here, she holds that "Lilburne and his colleagues were concerned with the equal distribution not of stuff but of standing," eradicating the distinction between peers and commoners and "claiming the rights and challenges of Peerage for themselves." (Bejan, 2022, 612). To gain leverage on the relative importance of social egalitarian arguments one of our tasks below is to assess the frequency and distribution of two linked adjectives that Bejan takes to be especially important: the conjunction of "equal and alike" and the disjunction of "peers or equals" (and variants thereof). Ultimately, we find these terms to not be especially common in Leveller thought and we provide a different perspective on the interpretation currently applied to them. We also argue that scholars who seek antecedents

---

[2]This term, "equal high standing", is used by Ober (2012) to characterize civic dignity, as Bejan notes.

for socially egalitarian, anti-hierarchical arguments ought to turn instead to the Diggers: we find that their understanding of "equality" is fundamentally different from that of the Levellers, and we attempt to characterize how it departs.

It is worth noting that contemporary egalitarians who turn to the Levellers for purposes other than an account of social equality tend to rely on the Rainsborough passage to support an anti-hierarchical vision of political agency. For instance, Jeremy Waldron cites this passage to emphasize the democratic value of "empowering those who would otherwise be powerless," suggesting that the central aim of a democratic constitution is to affirm the formally egalitarian political authority of the "poorest he," as opposed to a nondemocratic society, which tends to assign power to those who whom it comes "naturally or who... have been vested historically with power over a given society." (Waldron, 2016, 37) As we show, once we turn to the questions pertaining to the institutional realization of political power, it becomes clear that the aim is not to elevate the "poorest he" as such, even if it is to (re-)establish the lost legal rights of the ordinary free-born Englishman. We show that the real focus of the Levellers is on the assertion of a set of *legal* entitlements, held both by the body of citizens and by Parliament, that have been denied and usurped, and to which all Englishmen by right have claim. Furthermore, we show that this focus on law and legal entitlements—over and above the question of franchise reform Baker (2013); Macpherson (1962)—is what makes the particular Leveller commitment to electoral institutions distinctive.

How to characterize these legal rights, and their salience within Leveller thought, is contentious. There is a long, distinguished historiography of the Leveller relationship with common-law tradition, including contributions by Hill (1958), Skinner (1965), and Pocock (1957). Famous among those issues was the Leveller understanding of the "Norman Yoke": whether William of Normandy's assumption of the throne reflected a fundamental break with Anglo-Saxon norms or was continuous with them. Some, like Pocock, have emphasized Leveller appeals to natural right as a basis for such legal freedoms, arguing that the invocation of legal rights would be incoherent given the imposed and illegitimate nature of the common law as a result of the breach. By contrast, Seaberg (1981) argued that Lilburne's criticism of the common law was mostly in the service of

a procedural argument on behalf of local justice, rejecting reliance on Latin, on lawyers, and on central pleading, and, more generally, that *Magna Carta* provided a sound basis for the guarantee of liberties but had had its force wrongly diluted. On Seaberg's account, the Levellers sought to argue that the fundamental laws and liberties of the English had been defended repeatedly against the use of arbitrary power, most notably including trial by jury (Seaberg, 1981, 801), and indeed had been affirmed by William in his coronation oath.

In recent work, Rachel Foxley has defended a conception of Leveller thought that takes seriously the invocation of the legal liberties of "free-born Englishmen," particularly on the part of Lilburne. Foxley demonstrates how regularly Lilburne appealed to parliamentary authority and to statute, in support of the claim that the law provided a source for Englishmen's rights, though he also was prepared to rely on both common law and natural law argumentation to support the need for citizens to exercise these rights. Indeed, in Foxley's view, "Lilburne's language actively remodelled individual legal freedoms into a unified conception of a set of citizen rights. Law was the mediating term in this transformation: freedoms under law retained their names but came to signify the single and universal freedom guaranteed by the law." (Foxley, 2004, 857) Yet other Levellers resisted these appeals to positive law and to Magna Carta in particular, and so it is important to assess the extent to which the appeal to law was mostly a feature of Lilburne's idiosyncratic vision. As such, we rely on these text-as-data tools to see how the appeal to legal liberties, and citizens' equal claim to their exercise, developed over time, both within and outside of Leveller pamphlets. We provide evidence that corroborates Foxley's claim that Lilburne was, indeed, an innovator— and that his "freeborn" construction was picked up by others. But we also show that this was not a "once-and-for-all" shift, either among the Levellers or their non-Leveller contemporaries. We also demonstrate, statistically, that Leveller ideological commitments are consistent and coherent outside of Lilburne's obvious personal influence.

Whatever one's view of the extant literature on the Levellers, it is clear that claims about them are fundamentally *comparative* in nature. That is, assertions that they produced novel arguments, or that Lilburne differs from his colleagues, or that they changed their focus over time, all require

assessing one group of texts *relative* to another. Furthermore, these claims implicitly (and some-times explicitly) need us to measure what *meaning* different authors attach to different words and phrases. We now introduce data and methods for exactly these purposes.

3 Data: Leveller *Pamphlets* and Parliament

Our core data are the words of the Levellers and their contemporaries. Predominantly, these are printed *pamphlets* which were circulated to local populations. Our source for these pamphlets is the *Leveller Tracts* project (Hart and Kenyon, 2018), distributed on the web by the *Online Library of Liberty*.[3] We will describe the pamphlets in more detail below, but we note that they are all transcribed from images of the original printed works. As might be expected with a relatively ancient collection, the "result has been files with thousands of illegible words and characters." The compilers of the texts therefore insert specific characters, notes or marginalia to alert readers to this possibility. We simply strip out these problematic cases: for a reasonably large corpus, the occasional missing term is unlikely to be devastating for our inference. We also generally dispense with footnotes and endnotes where we find them. Ultimately, we have 260 texts from the seven "main" volumes, plus three addendum volumes.[4]. Table 1 provides an overview of their lengths: the mean is around 12000 words, but the median is just over 6000 words.

Min	25th percentile	Median	Mean	75th percentile	Max
196	3030	6272	12035	13490	111256

Table 1: Lengths of texts (numbers of tokens) in our "core" Leveller corpus.

Much of this skew is caused by three or four unusually long composite documents—including the *Putney Debates*, the *Whitehall Debates* and several consecutive issues of Lilburne and Over-ton's *The Moderate* newspaper—which we treat as single entries to the corpus here. Note that

some composite documents, like the *Putney Debates*, are essentially reports of conversations or discussions between actors. We do not consider these "Leveller" (or non-Leveller) pieces *per se* precisely because that contain other voices that are not part of a recognizable single movement.

Our core corpus begins with a document from March 1638, and ends with a document from July 1666. For orientation purposes, we think of the "main" period of Leveller activity as being that from August 1645 (Lilburne's first writing from imprisonment) until November 1649 (approximately when the authors of *The Moderate* newspaper were jailed). Within this 1645–1649 interval, we have the *Putney Debates*, the publication of the *Agreement of the People* manifestos, and the Banbury Mutiny—which arguably marks the last time that the Levellers were a force in the New Model Army. Where we do not know the precise date of a document, as given by the online library, we impute it.[5]

In terms of authors, specifically first authors listed, the data is also skewed. Leveller leaders, such as Lilburne, Overton, Walwyn and Thomas Prince, are responsible for around 40% of all tracts. To these "main" authors—who co-wrote *The Agreement of the People* manifestos from captivity—we add John Wildman and Thomas Sexby, as constituting an expanded version of the term Leveller. To get a sense of the balance between these groups and others overtime, consider Figure 1. There, the [red] circles and [pink] squares mark Leveller texts ("main" or from "other" authors), while the [blue] triangles are non-Leveller texts (by our definition). Notice that the *y*-axis is the log (base 10) length of text, and thus the Leveller and non-Leveller texts are somewhat balanced on this dimension over time. Our "main" period—where most of the data appears—is demarcated by broken bold lines at the center of the plot.

---

[5]If we know the month, but not the day (about 8% of the data), we take it as the 15th of the month; if we do not know the month (about 5% of the data), we take it as July.
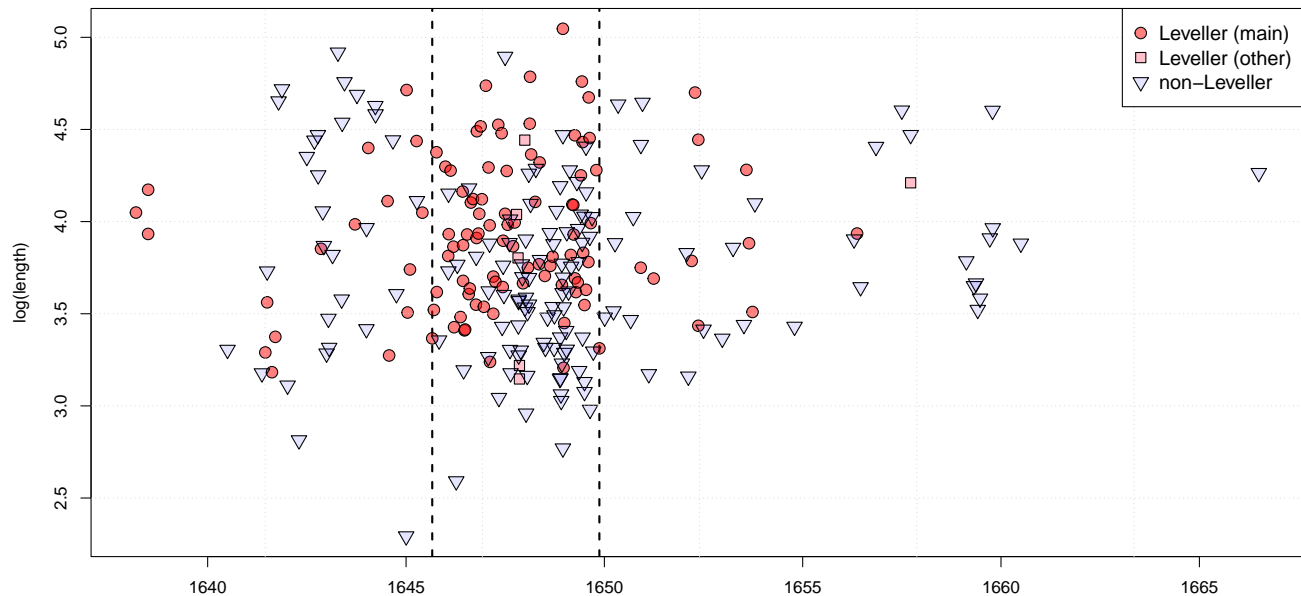
Figure 1: Authorship of core corpus, overtime and by (log) length. The two vertical dashed lines signify our "main" period of study

3.1 The Challenge of *Early Modern English*

The texts we will work with are all written in *Early Modern English* (EME), a form of the language used between the Tudor period and the Restoration (1660). Though certainly intelligible for modern readers, EME is not always trivial to understand. Partly this is a matter of word use and choice, and partly it is grammatical. For one thing, spellings are inconsistent. For another, certain terms which are unused now (such as "thou" as a second-person pronoun) are very common. And there is a generally bombastic tone in addressing readers. Consider this snippet from *The Grand Concernments of England ensured*, published in October 1659 (capitalisation as original):

> IF thou art prejudicate, save thy purse and thy paines; 'tis the considerate man, he that ponders his wayes, I had rather deal with: I promise thee thus much, I have no design to seduce thee, but whatsoever I have written, is my very thought: it may be thou art perswaded thou maist better employ thy time in Reading; I believe no lesse, however thou shalt finde some things here, not altogether unworthy of Consideration.

We might "translate" this as something like

variant	use case	count
free-borne	free-borne people	1
Free-born	Free-born men of England	5
free Birth-right	Birth-right and Inheritance	1
free Denizons	free Denizons of this Realm	1
Freemen	Freemen of London	5
Free-men	Free-men of London	17
freeman	freeman of this Kingdom	2
free-man	every free-man therof ought	7
Free mans	particular individuall Free mans	1

Table 2: Variants of "freemen" in one document.

If you are prejudiced, don't bother reading. This is for considerate, thoughtful people. What follows is not my attempt to bamboozle you, but just my genuine thoughts on these issues. I for one think my words are worth reading, and hopefully you will agree if you press on.

Obviously, some care is required in dealing with such texts. In particular, off-the-shelf methods trained on modern documents are unlikely to perform well, especially if they mistake unusual (by our standards) spellings for (uniquely) rare words. More generally, we need other texts of the period to contextualize these ones. We discuss some mitigation strategies below.

For now, we note that simply tracing the use of a term like "free-born subject" through time is difficult for EME. To see this point consider a single document in our collection, Lilburne's *London's Liberty In Chains discovered* (1646). It uses terms similar to `freemen` at least nine ways. In Table 2 we give those nine varieties some example uses (to establish that they are approximate synonyms) and their total counts. We will return to the nature of this challenge below.

4   Methods: Key words and Embeddings

Aside from tracing particular terms, our questions concern the nature of meaning, and how those meaning differ across groups of people. We will require a way to plausibly measure meaning, and understand how it varies by covariate values (the *X*s in a regression).

As regards the first challenge, recent work in social science has turned to "word embeddings" (see Rodriguez and Spirling, 2022, for an overview). These approaches use neural networks to model the probabilities of words occurring in sequences, such as they would appear in sentences. Ultimately, they yield real valued vectors—of say, 300 length—that consist of estimated weights from the probability models that are attempting to accurately predict which words will follow others, after observing many millions of such sequences. At a higher level, these models rely on what is known as the "Distributional Hypothesis" that one will "know a word by the company it keeps" (Firth, 1957, 195). There are two related parts to this insight. First, words that appear near (physically around) other words are likely to be connected in terms of meaning. For example, polemicists in a given period may regularly write of the "right to vote". An embedding model might then learn that the "vote" is something that one might have a "right" to. Second, suppose that we also observe the phrase "suffrage right" in the corpus, but that this term is never used by those who speak of the "right to vote". Nonetheless, via the connection to "right", the model will learn that "vote" and "suffrage" are related (here almost synonyms). In practice, this means that the estimated vectors for "vote" and "suffrage" are similar, in terms of the size and sign of the numbers those vectors contain. By extension, the model will also learn that certain words have opposite meanings (their vectors are negatively correlated) or not really related (the correlations are zero).

Social scientists have realized the potential of such ideas, and applied them to subjects as diverse as parliaments (e.g. Rheault and Cochrane, 2020) and cultural norms (e.g. Caliskan, Bryson and Narayanan, 2017). In all cases, researchers must make decisions about what underlying corpus the embeddings should be fit to. Generally speaking, that corpus should be as large as possible, but it should also reflect the ways terms are used for the subjects under study. For our period of EME, this is a challenge. One option would be to simply ignore the subtleties of the time and place we are studying, and use, say, modern embeddings (pre)trained on Wikipedia or similar corpora. This is unlikely to be successful. Another option would be to avoid this implicit (and likely bad) approximation, and simply fit all our embeddings to the Leveller corpus itself. But that corpus is

small: it is just 260 documents (two million words). By way of contrast, the underlying corpora for readily available pretrained embeddings for models such as `GloVe` (Pennington, Socher and Manning, 2014) have token numbers in the *billions*. In our corpus this means that, for example, a term like "elective" occurs fewer than 60 times. Simply put, this will not allow for precise or useful embeddings.

We resolve this problem by finding a source of EME to help train our model. From Rodon and Paskhalis (2021) we obtained written proceedings of the English parliament for the relevant period. In particular, we have the "journals" of the House of Commons and House of Lords, starting in 1547 and 1510 respectively. These are non-verbatim summaries of daily proceedings. And we have the statutes passed by parliament from 1625 onwards, including the interregnum period. All told, this is some 19 thousand documents, and some 35 million tokens.

With this relatively large corpus available for pretraining, the next problem is methodological. We are required to bring this data to bear on our relative small set of tokens for the focussed Leveller period. Here, we use ideas laid out in Rodriguez, Spirling and Stewart (2021*a*). Those authors provide a particular implementation of "a la carte" embeddings (Khodak et al., 2018). The essence is that one can form a high quality embedding representation for a given word, by taking the (reweighted) average of the embeddings of words *around it*. This is helpful for situations, as here, where the words may not be especially frequent—either because the token of interest is rare, or the corpus is quite small, or both. But because we have a relative large period-specific corpus, we can build up period-specific representations. In this way, the reweighted "a la carte" averages we take will tend to be high quality, insofar as they correctly capture how given terms were used at the time in question. Rodriguez, Spirling and Stewart (2021*a*) place this technique in an underlying statistical framework that allows for regression-style work. In short, they allow each instance of a term in a corpus to have its own embedding, and this matrix of embeddings to then be regressed on covariates. Specifically, this is a multivariate regression (commonly used when each element of the outcome is a vector). The authors use non-parametric resampling techniques to allow for hypothesis tests. One can test whether the embeddings for one group (say, Levellers prior to the

Civil War) differ in a statistically significant way to another group (say, Levellers after the Civil War). Beyond this, Rodriguez, Spirling and Stewart (2021*a*) also allow for detailed description of *how* usage differs by reporting nearest neighbors (literally, the words closest in embedding space) for each group.[6]

Implicit in our approach is the notion that the way actors *use* language is indicative of what they *mean* by that language. And thus, to the extent that the former differs across groups, so does the latter. Obviously, there are many ways in which this may not be true, and there is a large philosophical literature demarcating more principled approaches (e.g. Schiffer, 1972). For one, actors are strategic, and choose terms they believe an audience will respond to in particular ways, or simply to differentiate their ideas from some other author. In this sense, there may be no "sincere" belief of a difference in meaning. Second, even if all statements are sincere, the contexts in which certain words appear reflect quantities beyond meaning *per se*. For example, Democrats and Republicans in Congress might discuss "immigrants" in radically different ways, but they do not literally disagree on the meaning of the term. Nonetheless, as a first approximation we believe our approach is reasonable and inline with other efforts in corpus linguistics (see, e.g., Camacho-Collados and Pilehvar, 2018, for an overview).

4.1 Validating the Embeddings

We can gain some immediate insight into the merits of our localized embedding strategy by exploring how it represents key words from the period. We can do this both in an absolute way ("does this meaning make sense?") and in a relative way ("does this meaning improve on other techniques?"). Table 3 displays some results for six words of specific importance in our period. In the "Standard GloVe" column, we report the (nine) nearest neighbors in embedding space for a pretrained standard GloVe model fit to modern English sources such as Wikipedia (available from Pennington, Socher and Manning, 2014). In the final column is our model, fit to the parliamentary

---

[6]Those authors have released accompanying software for all these tasks (see Rodriguez, Spirling and Stewart, 2021*b*).

token	Standard GloVe	Our Model
sovereign	sovereignty, entity, debt, status, downgrade, creditworthiness, entities, affirmed, monarch	reign, gracious, loyal, subjects, dutiful, majesty's, memory, sacred, queen
parliament	parliamentary, mps, assembly, legislature, duma, lawmakers, deputies, elected, speaker	assembled, lords, present, commons, now, whereas, hath, may, house
cromwell	1649, 1653, oliver, 1658, 1651, 1654, 1652, 1655, 1657	ireton, oliver, maynard, gen, north, lieutenant, grey, herbert, wentworth
rights	human, freedom, freedoms, liberties, abuses, advocacy, advocates, activist, legal	liberties, just, priviledges, ancient, freedoms, subject, privileges, preserve, freedom
alms	zakat, begging, gt2, lmp2, gt1, prayers, mans, scavenging, blessings	lot, housekeepers, receiving, voted, freeholder, elected, housholders, majority, poll
god	allah, gods, divine, heaven, christ, bless, jesus, faith, worship	world, even, heart, people, gods, almighty, spirit, glory, things

Table 3: Nearest Neighbors (9), Standard GloVe v our locally trained model.

data and the various Leveller texts noted above.

We will not belabor this point, but from inspection it is obvious that our model returns terms that are more meaningful—and thus accurate—for our corpus of study. For example, the term sovereign is associated with being a "gracious", "dutiful" "reign". For the standard model, it is associated in part with "debt" and a potential "downgrade". Similarly, rights in our case returns associated terms of the age like "ancient" "priviledges". For the standard model, we see modern terms such as "advocacy" and "human [rights]". The word alms may seem obscure, but provides final evidence of the quality of our local model. In particular, "alms" or rather, the receiving of alms, were a potential disqualifier for Leveller notions of suffrage. Or model correctly locates this terms near "freeholders" and "housekeepers" who generally were thought worthy of the vote. The standard model meanwhile returns "zakat", an Islamic term for almsgiving, and "gt2", a type of BMW car. Having provided some basic validation, we now move to discussing our tests and results.

5 Results

We begin our results by studying the "keyword" claims for Lilburne and the Levellers more generally, as espoused by Foxley (2004) and Bejan (2022). We then move to the broader issues of "meaning" and our embeddings models of the same.

5.1 Lilburne's Innovations: "freeborn Englishmen"

To answer our first set of questions, we subset our pamphlet data to the (extended group of) Levellers, with special focus on Lilburne. We use regular expressions, including wildcard characters to track the various uses of `freeborn` and its allies, over time.[7]. As usual with such pattern matching problems, there is a danger of both false positives (we identify a phrase which has a meaning different to that intended) and false negatives (we miss some expressions which we should have included). We guard against the latter this with some manual checking of context; we guard against the former by loosening the match requirements and seeing if this radically increased the number of "hits" (it did not).

As regards Lilburne, two findings emerge. First, to the extent that he transitioned from using "freeborn subjects" to "freeborn Englishmen", it was not once and for all. In particular, Lilburne's first use of any "freeborn" type term is a kind of halfway house; he uses `freeborne English Subjects` in January 1645 (*A Copie of a Letter* ostensibly sent to William Prinne or Prynne, the Puritan polemicist). The specific context is Lilburne's recounting of recent history and his belief about the purpose of parliament. However, this does not mean Lilburne never uses the `freeborn subject` construction. Interestingly, it is only *after* 1645 that he uses `freeborn subjects` at all: in *Vox Plebis* (November, 1646) and *The Charters of London* (December, 1646). After this, he does not use the `subjects` construction, and moves wholly to using the `English` version (so, 1645 to 1653). Second, to the extent that we believe Lilburne is an innovator, we should expect to see that—even though he does not make a clean break of usage in his own work—he is nonetheless

---

[7]Specifically the variants are: `freeborn *`, `free-born *`, `free born *`, `freeborne *`, `free-borne *`, `free borne *`, `free-man of *`, `free-men of *`

the first mover. This is true. Specifically, Lilburne uses a "free-man of England" construction to refer to himself in October 1645 (*Englands Birth-Right Justified Against all Arbitrary Usurpation, whether Regall or Parliamentary, or under what Vizor soever*). The next use of the term is by Walwyn a few days later—to refer to Lilburne himself (*Englands Lamentable Slaverie Proceeding from the Arbitrarie will*). After this, the "Englishman" construction turns up some 29 times, of which of which 21 have Lilburne as first author.

As we stated, a feature of our comprehensive corpus is that we can compare the Levellers to their contemporaries. For example, it could be that Lilburne picked up the "subject" to "Englishman" innovation from someone not part of the usual Leveller set. We do not find that to be the case. When we look at who uses the "subject" construction in our corpus (ignoring anonymous works), it is authors like Thomas Johnson (1645, 1646), John Selden (1646) and Willian Prynne (in 1653). The first of these uses (April 1645) is *after* Lilburne's first use. As for the "English" construction, this turns up in documents by William Larner (coauthored with Lilburne, 1646), Robert Lockier (1649), William Prynne (1648, 1649, 1653), James Freize (1659) and Robert Bennet (1649). All of these incidences are *after* Lilburne's first use. So, to the extent that we believe Lilburne influenced the other Levellers, it seems plausible he directly or indirectly influenced other thinkers of the time too.

5.2   Other phrases: "equall" and "equals"

Our second set of phrase-based queries concern the use of terms deemed by earlier scholars to be indicative of broad attention to matters of social equality. To recall, these are "equall and alike" and "peers or (and) equals". Our immediate observation is that neither of these terms is very common in the corpus. In particular, "equall and alike" occurs just three times.[8] Twice, the term is used by Lilburne (*The Free-mans Freedom Vindicated* (1646), *Regall Tyrannie discovered* (1647)) and once by Overton (*An Appeale from the degenerate Representative Body* (1647)). It is true that the contexts of these terms concerns notions of "by nature" and "natural birth", but we are cautious

---

[8]For completeness, we note that "equall or alike" never occurs.

about updating too much for such a low number of uses. A similar story arises for "peers or (and) equals". It is used just six times by the Levellers, and only ever by Lilburne in four documents (1646, 1648, 1649, 1652). In addition, the context here does not seem to connote social equality *per se*. In five of the cases, Lilburne is referring to jury trials. In the fifth usage, he makes the point that the members of the House of Lords are not a jury for the purposes of imprisoning freemen (including himself). That is, from *The Laws Funerall* (1648): "the Lords are none of my Peers or Equals, and therefore are none of my legal Judges, nor have not the least Jurisdiction (in any case whatsoever) in the world over me."

In sum, while we agree that Lilburne was an innovator with respect to his "freeborn" construction, we are skeptical of extant claims about other language suggesting a deep commitment to *social* equality rather than legal equality.

5.3   The Diggers and the Levellers

As a final part of our *social* equality investigation, we now compare Leveller and Digger thought.[9] If we treat the latter group as they regarded themselves—"True Levellers" committed to a new non-hierarchical social order—then we have a potential fixed point against which to measure Leveller commitments. To do this, we turn to the embeddings methods we introduced above. In particular, we compare Digger and Leveller texts as regards their use of words that connote "equality".[10] Candidly, we reiterate that we do not have much data on the Diggers: only four documents, and only one use of an equality term, so readers should be cautious about over-interpreting the results. Still, this is a comparison we are statistically able to do.

Our first cut on the problem is to compare the two groups in terms of their nearest neighbors for (the combined) "equality" term. Here, we are asking which embedded terms in the corpus are

---

[9]By "Digger thought" we mean the four documents by Gerald Winstanley in our corpus: *A Watch-word to the City of London* (1649), *The True Levellers Standard Advanced* (1649), *An Appeale to all Englishmen* (1650) and *An Humble Request* (1650).

[10]We used regular expression matching for 'equal' and 'equitable' resulting in a dictionary of: "equal", "equall", "equallity", "equalls", "equally", "equals", "unequall", "equitable". Of course, "equity" itself also has a technical legal meaning as the form of justice rendered by the Court of Chancery. The term "equitable" in the corpus generally does not reflect this meaning, but instead is synonymous with "equal" or "fair."

closest—via cosine similarity, a standard measure—to each group's understanding of the focus term (here, "equality" in its various guises). The results are in Table 4. We see that the Levellers use "equality" in a way closest in connotation to the use of "lawes", "freedome" and "selves" in the pamphlets. Meanwhile, the Diggers use "equality" in a way closest to "mankind", "righteous" and 'righteousness".

Table 4: Levellers and Diggers have different connotations for "equality" (10 nearest neighbors)

rank	Levellers	Diggers
1	lawes	mankind
2	freedome	righteous
3	selves	righteousnesse
4	equall	righteousness
5	legall	earth
6	fundamentall	glory
7	freedomes	freedom
8	peoples	darknesse
9	self	spirit
10	common-wealth	soul

While there are some differences, there are also some similarities—e.g. both groups have versions of "freedom" as nearest neighbors. To throw the comparison into starker relief, we now report the nearest neighbor terms that are *most distinctive* to each group. The results are shown in Figure 2. To understand that figure, and some others that follow, note first that for a given group (say, the Diggers) we can provide an embedding for any term (say, "equality") and then calculate the cosine similarity of that term to any other (say, "mankind"). We can do that for another group too (say, the Levellers), and thus we have two cosine similarities between two terms, for each group. The *ratio* of those two cosine similarities can take various values: if it is 1, then the cosine similarities are both the same (top and bottom), we draw the conclusion that both group use the term identically (at least relative to the focus word). If that ratio departs from one, it means that either the group in the numerator has a much larger cosine similarity than that in the denominator, or that the group in the denominator has a much larger cosine similarity for the terms than the group in the numerator. The figure shows the (top) terms that have (absolute) cosine ratios that

are distinct from one. The *x*-axis is simple the (raw) value of that ratio. The *y*-axis is the ratio's absolute deviation from 1, and makes comparisons of magnitude a little easier.

On the left of the figure then, we have the terms for which the ratio is lower than one: that is, the terms for which the Diggers have higher similarities to "equality" than the Levellers. To the right of the plot are the terms for which the Levellers have higher similarities to "equality" than the Diggers. Terms nearer the top of the plot are words for which the cosine ratio magnitudes is most different from one—so, terms used in ways that are most distinct to the two groups (Diggers and Levellers). The plotting characters at the bottom provide information about where the words themselves exist in terms of being nearest neighbors in the sense presented in Table 4: some of the most distinct terms are nearest neighbors for the Diggers use of "equality" (e.g. "mankind", marked with a circle) , some of the most distinct terms are nearest neighbors for the Levellers (e.g. "selves", marked with a square), and some in both lists (e.g. "freedome", marked with a triangle). The terms with statistically significant cosine ratios (for which a ratio of one is the null hypothesis) are marked with asterisks.

Examining Figure 2, we see that—to the far right and top—the most distinctive term that it is statistically significant is "legall". Equivalently, the Levellers use the term "legall" in the close context of "equality" in a way the Diggers do not. This is *prima facie* statistical evidence that Levellers and Diggers have different understanding of "equality" and the Leveller one is legal, in nature.

Finally, we run an embedding regression of all the various "equality" terms on group member-ship (Leveller authorship as a binary variable) for this sub-corpus. The results are reported in Table 5. The coefficient is on the dummy variable `Leveller`, which simply connotes whether the author of the usage of the term in question was a Leveller (1) or not (0). We give the coefficient itself (3.029) and its estimated standard error (0.435). We also obtain an empirical *p*-value, which here is less than 0.01. The substantive point is that the Diggers and Levellers differ in their contextual use of these terms.

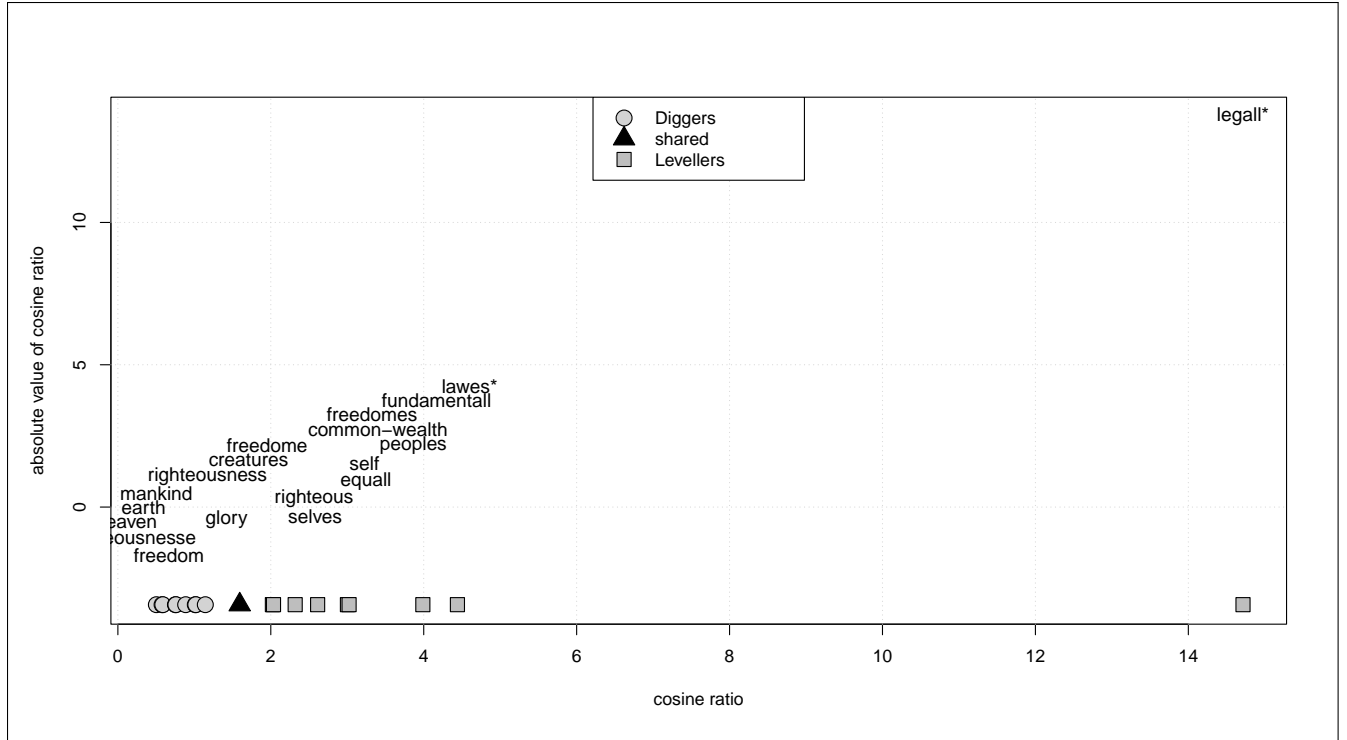Our conclusion here then is that *if* we regard the Diggers as exemplars of social equality, then

Figure 2: Digger v Leveller most distinctive nearest neighbors of "equality": Levellers 'legall' understanding is different to that of the Diggers. Asterisk indicates $p < 0.01$.

	equity
Leveller	3.029***
	[0.435]

Table 5: Levellers and Diggers use the term 'equality' (and its variants) in a way that is statistically significantly different ($p < 0.01$).

the Levellers are not that. But exactly what the Levellers *did* mean by "equality" is still unclear—

our next set of results speaks to this.

5.4   What the Levellers Meant by "equality"

We begin this part of the analysis by comparing the Levellers to *all other* authors, in terms of their understanding and use of 'equality'.[11] First, as with the Leveller-Digger comparison, we report the ten nearest neighbors by cosine distance.

---

[11]Note that "all other authors" excludes the Diggers, and a set of anonymous documents. We use the same dictionary for "equality" as above.

Table 6: Levellers have different connotations for 'equality' (10 nearest neighbors) relative to their contemporaries.

rank	Levellers	non-Levellers
1	lawes	peoples
2	freedome	absolute
3	selves	common-wealth
4	equall	equall
5	legall	publike
6	fundamentall	freedome
7	peoples	lawes
8	freedomes	selves
9	self	indeed
10	common-wealth	must

Once again, we notice the Leveller preoccupation with "equality" as a legal concept—"legall" and "lawes" appear high in the rank ordering of terms. For non-Levellers, we see "absolute", "peoples" and "publike" before we note "lawes". As above, we can refine this assessment by studying the *most discriminating* nearest neighbors, in terms of their cosine ratio. Figure 3 does this. Note that, in line with Table 6, the terms "legall" and "law" are indeed distinctive to the Leveller understanding of equality; meanwhile, non-Levellers are distinct in that they use "publike", "people's" and "common-wealth" (among other terms) when discussing equality.

Descriptively then, we have some evidence of Leveller difference—and, again, it appears to focus on 'legal' (or "legall") nature of equality. We can provide statistical corroboration of this claim by running a series of regressions comparing the embeddings for various terms, across the groups. We do this in Table 7. The two sets of columns represent regressions for the terms "equall" and "equall", "equalls" and "equally", together, respectively. We denote the second column as "equall plus" to make the point that it is comparing more than one term. Within those two groups, we include a specification with and without a time fixed effect. Specifically, we include a dummy for the period after the publication of the final *Agreement of the People* in May 1649. The idea of this fixed effect is to ensure that our basic results are not simply an artefact of a given historical period in which the Levellers were very active. As can be seen, for every specification, Levellers
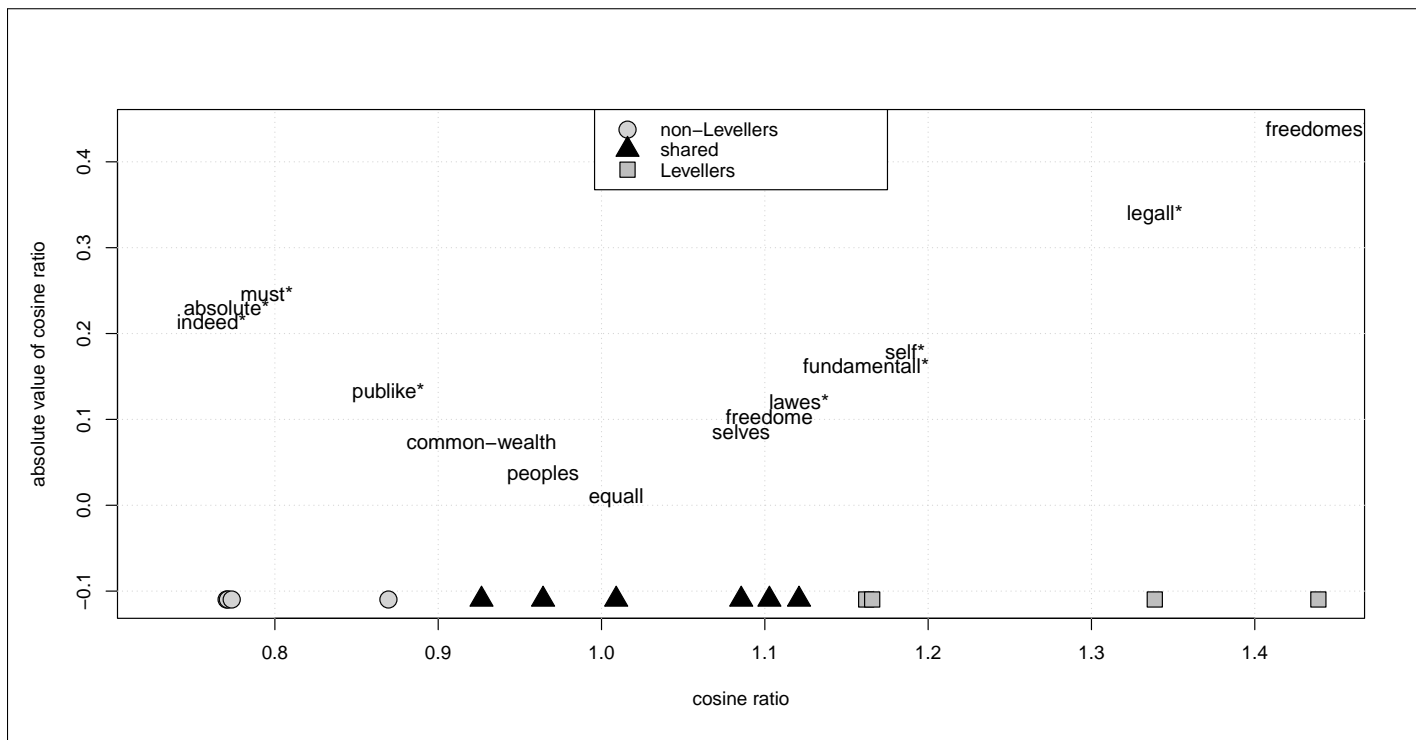
Figure 3: Non-Leveller v Leveller most distinctive nearest neighbors of 'equality': Levellers 'legall' understanding is different to that of non-Leveller contemporaries. Asterisk indicates $p < 0.01$.

and non-Levellers use these terms differently.

	Terms			
	equall	equall	equall plus	equall plus
Leveller	0.961***	0.954***	1.032***	1.045***
	[0.062]	[0.067]	[0.065]	[0.063]
Time Effect	✗	✓	✗	✓

Table 7: Levellers use the term 'equall' and related terms in a way that is statistically significantly different from non-Leveller contemporaries($p < 0.01$). Time Effect is a dummy variable denoting whether the document in which the term appears was published before or after May 1649.

A reasonable concern at this point is that the differences we observe on "equality", while not due to time effects, might be due to *author* effects. Put otherwise: perhaps the differences between Levellers and non-Levellers are driven purely by Lilburne's presence in the former set. We know from our results above that he innovated on particular terms, and it seems possible that the variants of "equal" might be included here.

Our findings are subtle: while it is true that Lilburne does differ from his Leveller contemporaries, he is not the sole factor behind the broader Leveller versus non-Leveller divergence. To see the former point, consider Figure 4. There, we report the terms that are most distinctive in the contexts of Lilburne's use of "equality" relative to his Leveller colleagues. For Lilburne, these are (*inter alia*) "lawes" and "legall".
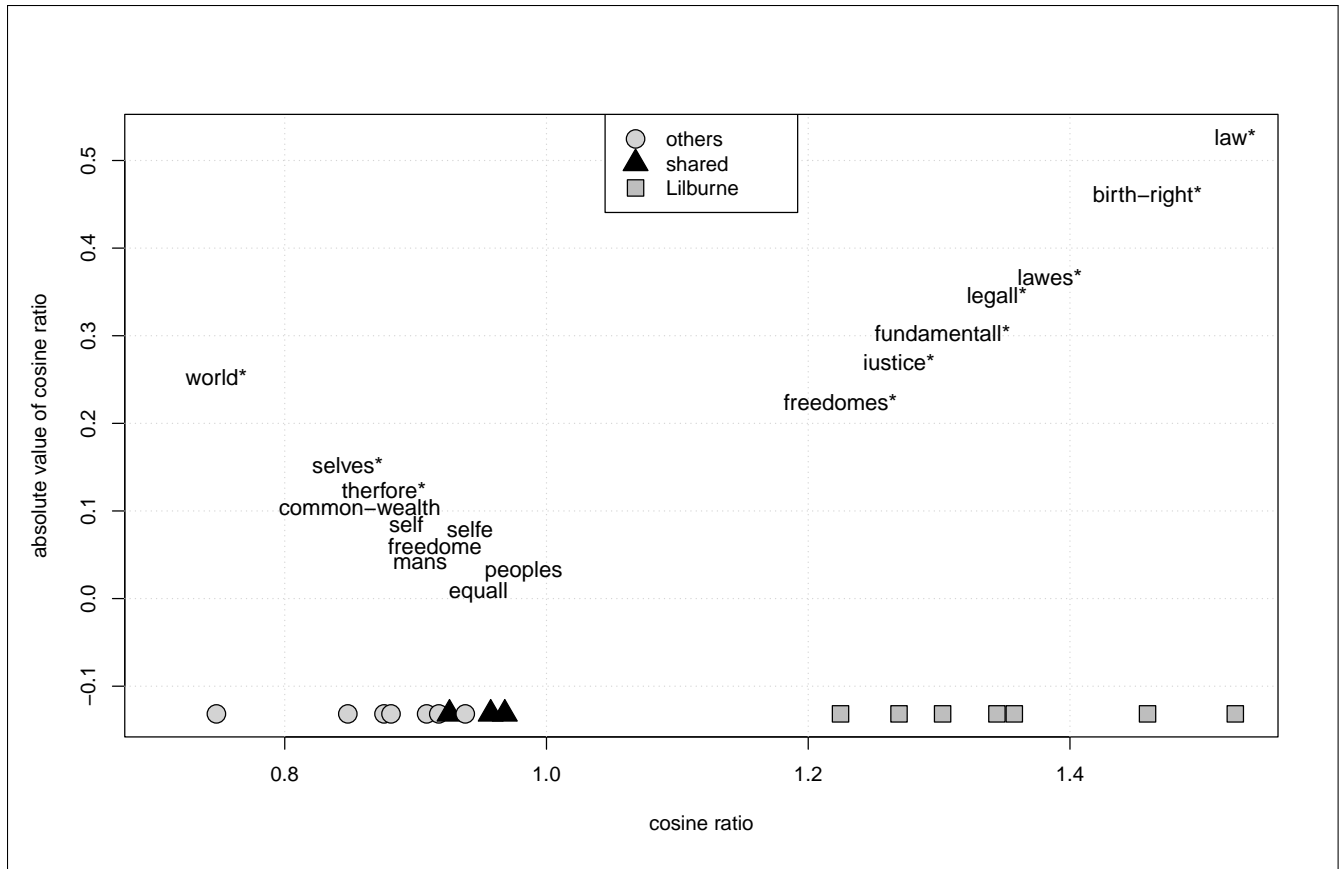
Figure 4: Lilburne differs from other Levellers in terms of his use of "equality". Asterisk indicates $p < 0.01$.

Importantly though, the more general pattern we observed above is not a mere Lilburne-as-author effect. We can see this from Table 8, where we re-run the core regressions from Table 7, adding either a dummy for Lilburne (on a per use of term basis), or we removed him from the data altogether. Neither change affects the basic fact that Levellers differ from non-Levellers in their understanding of "equality". That is, that variable is always significant in every specification.

	Terms			
	equall	equall	equall plus	equall plus
Leveller	1.0746***	1.071***	1.084***	1.012***
	[0.085]	[0.077]	[0.061]	[0.068]
Lilburne effect	author FE	removed	author FE	removed

Table 8: The difference between Leveller and non-Leveller understandings of "equality" is not (merely) a consequence of Lilburne authorship.

A final comment in this subsection concerns the use of "legall" in the context of "equality". We have shown that it is distinctive to Lilburne *and* distinctive to the Levellers. But conceivably, this might simply because Lilburne and the Levellers use it more often, perhaps because they discuss topics that require it more frequently. Fortunately for our general story, this is not the case. In particular, "legall" is used (at least once) in 80 documents in the subcorpus; 50 of those documents mentioning "legall" are authored by Levellers (or 0.625 as a proportion). This compares to 113 Leveller-authored documents of 200 total documents in the subcorpus (or 0.565 as a proportion). A simple proportions test suggest that Leveller authorship is not statistically significantly more likely among the documents mentioning "legall" ($p = 0.061$). Put otherwise, there is not immediate reason to believe that the difference in the context of Leveller discussions of "equality" is driven merely by topical choices.

5.5 Context-in-context: qualitative evidence on "equality"

It is reasonable that readers may want to see more qualitative information regarding what are, explicitly, claims about differing *contexts*. Here, we look in more depth at documents which contain proto-typical examples of the use of "equality" in context. Specifically, for a set of embeddings of the term and a set of contexts (literally words around our words of interest), we inspect the $N = 20$ nearest documents that contain those contexts. On the Leveller side, we study four of

27

these: three by Lilburne and one by Overton.[12] On the non-Leveller side, they are from various authors, including Rushworth, Saunders and Filmer.[13]

The preoccupation of four Leveller essays is the purported usurping of power by the House of Lords. In Lilburne's work, the claim is that the House of Lords has no right to try him because they are not his equals. In *Out-Cryes* and *Regall Tyrannie*, he cites Magna Charta and Lord Chief Justice Edward Coke's interpretation in favor of this position. In *Out-Cryes*, he distinguishes between the "PARES" of nobility and the "Commons of the Realms." Within each category, everyone is a "peer to another, even if he be of a several degree." Likewise, in *Regall Tyrannie*, Lilburne argues that commoners can only be judged by their "peers," who Coke "saith, is meant equals."

Lilburne repeatedly invokes the claim that no man shall be dispossessed unless it be by the "verdict and judgment of his Peers, that is equals." He characterizes the demand in favor of an "open" or "public" trial as one for a "fair and equall tryall." *Regall Tyrannie* does also argue for legal equality more generally: that Englishmen are born equally, and further that the law of the land is both an "equal inheritance" and grants them an entitlement to equal freedom under the laws of the land. But in *Whip*, the sole use of the term "equalls" is also with respect to the jury: Lilburne is arguing that if he is charged, he should be tried according to the common law before an "ordinary judge, the true and proper executor of the law" and a jury of his "Equalls."

Lest this seem to be only a function of Lilburne's interest in evading trial by Lords, it is worth noting that Overton's use of equality also primarily pertains to the jury. Although Overton begins by repeatedly invoking the claim that all men are born "equal and alike" by "natural birth," and that everyone aims at freedom from subjection to one's neighbors and that all recognize it to be "most reasonable, equal, and just" (note here that "equal" here stands in for what is elsewhere occupied by "equitable"), the focus quickly shifts to the wrongdoing of the Lords. The people are subject to the "incompetent, illegal, unequal improper judges," at odds with the passages from Magna Carta cited by Lilburne.

---

[12]Specifically: *The out-cries of Opressed Commons* (1646); *Regall Tyrannie discovered* (1647); *A Whip for the present House of Lords* (1648); *An Arrow against all Tyrants and Tyranny* (Overton, 1646)

[13]Specifically: *The Humble Petition of Several Colonels* (1654), *A Petition concerning the Draught of an Agreement of the People* (1649), *The Anarchy of a Limited or Mixed Monarchy* (1648).

The only other notable and distinctive use of the term "equall" comes in *Out-Cryes*, pertaining to the "equall power" of all the members of Parliament in their challenge against the King. So even here they do not use the language of equality for the aims of standing as peers; rather, their equality is simply with respect to power as MPs, and the objection is that the law is designed to secure them as unequals.

Non-Levellers tend to use the term "equal" in this latter fashion, primarily pertaining to the equal distribution of power, both in the distribution of representatives among the people and among houses of parliament and the monarchy. This latter comes in a long response from Robert Filmer to Philip Hunton, where he also swiftly rejects the putative "equall freedome from subjection" to consent to a monarch). The 1649 Petition from Rushworth focuses on rejecting the unequal distribution of seats (as against rotten boroughs), and that the people ought to be "indifferently and equally" provided for, perhaps along the lines of the "equitable" language we see elsewhere. The later, 1654 petition by Thomas Saunders also focuses on representation, and the need for Parliament to be "equally chosen by the People." All told then, our qualitative assessment agrees with our quantitative one: Levellers use "equality" in a particularly legalistic manner, and in a way dissimilar to non-Leveller contemporaries.

5.6   Electoral Institutions

Our final result concern claims about Levellers interest (or lack thereof) in the institutions of democracy. Here again, we compare Leveller authors to their contemporaries, but this time with respect to the use of terms connected to elections.[14] In Table 9 we report the relevant embedding regression, and it is clear that for the specific term "election" at least (mentioned 321 times in the subcorpus), Levellers differ from contemporaries in terms of their use patterns.

---

[14]In particular, these were: "elect", "elected", "electing", "election", "elections", "elective", "elector", "electors".

	equity
Leveller	0.995***
	[0.056]

Table 9: Levellers and non-Levellers use the term 'election' in a way that is statistically significantly different ($p < 0.01$).

But this regression result belies a more nuanced story relative to that around "equality". To see this, consider Figure 5 where we report the most discriminating terms with respect to cosine ratios for Levellers and non-Levellers on this subject. Interestingly while words like "lawes" and "chuse" and "people" are statistically significantly closer in context to the election terms for the Levellers, these are nonetheless *shared* nearest neighbors. Put otherwise: with some exceptions (like "freedome"), the Levellers and the non-Levellers understand elections in generally similar terms, even if it is the case that the distances differ slightly.
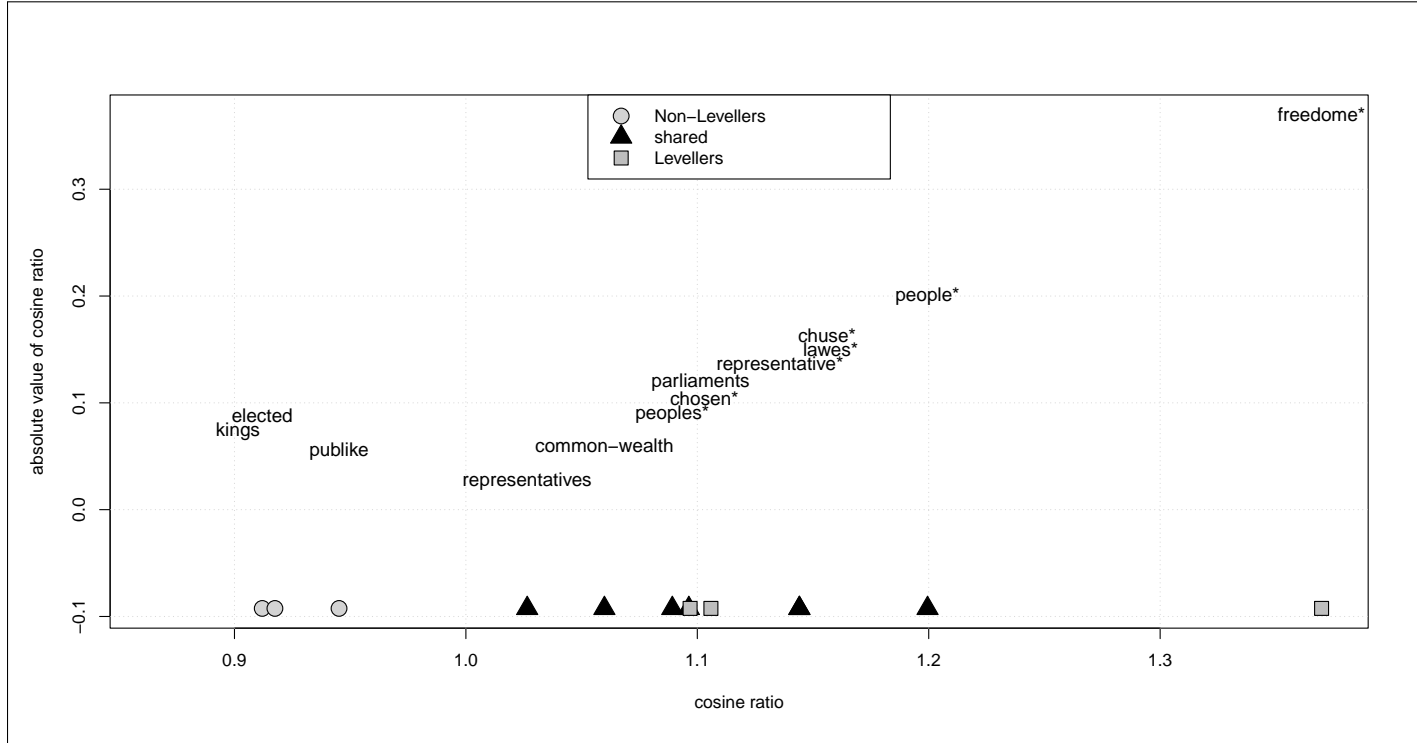

Figure 5: Levellers differ from non-Levellers on "election" context, but share several terms. Asterisk indicates $p < 0.01$.

One possible conclusion is that, fundamentally, the Levellers and their contemporaries might be more similar in their discussion of elections than they are in their discussions of (legal) equality. To get at this question more directly, consider Figure 6. There we fix the mean Leveller embedding for the equality (top) and election (bottom) terms respectively. Then, for each document in which these terms appear for non-Levellers, we report the cosine distance between that non-Leveller embedding and the average Leveller one (as a [red] circle character). If this distance is 0, the non-Levellers are discussing the terms in completely different ways. If this distance is 1, the non-Levellers are discussing the terms identically. We impose a simple loess so that the overall patterns are more obvious (unbroken black line); to orientate the reader, we also include a broken vertical line on the date of the final (failed) *Agreement of the People* (May, 1649).
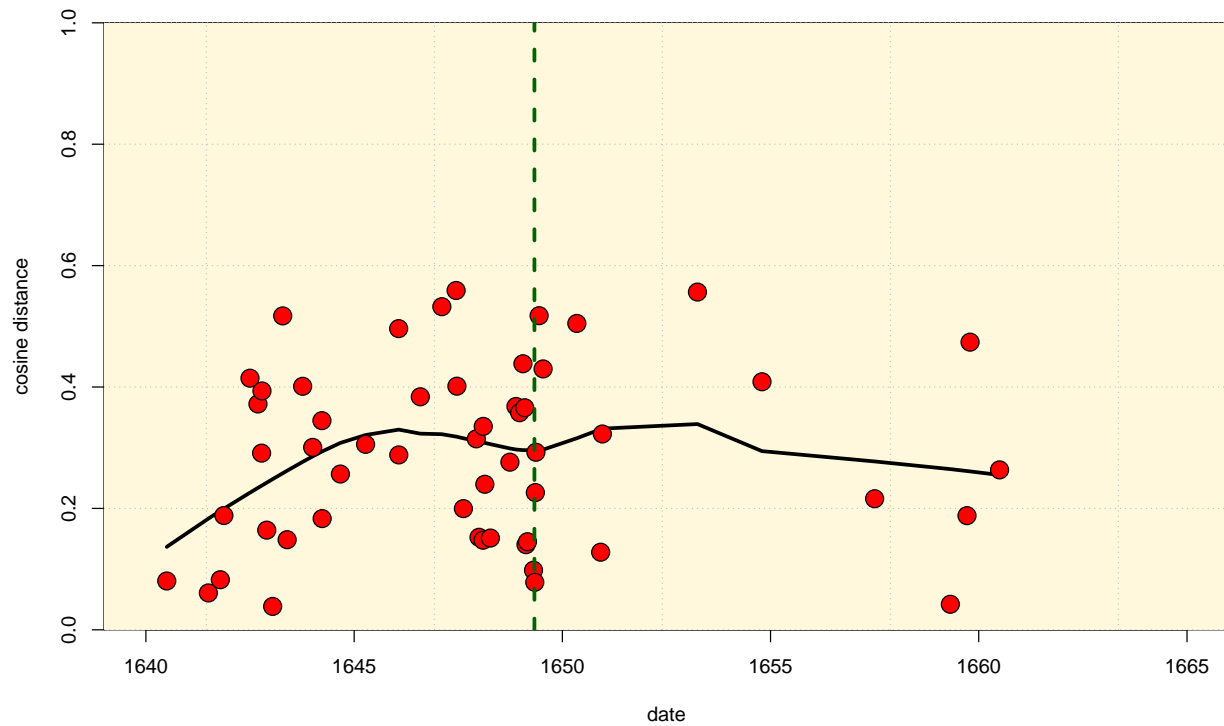
(a) Equality: Levellers v Non-Levellers over time.
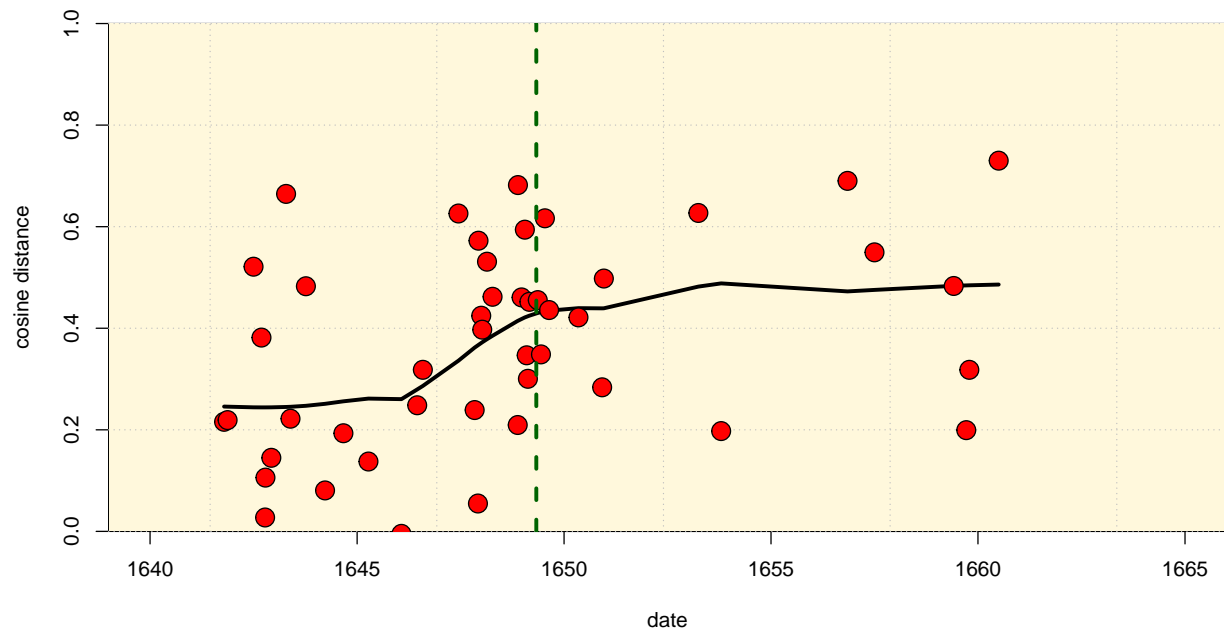
(b) Electoral Institutions: Levellers v Non-Levellers over time.

Figure 6: Comparing distances: Levellers and their contemporaries on equality and elections, over time. Each plotting point is the cosine distance between a non-Leveller document use of a term, and the Leveller mean embedding of that term. The unbroken black line is a loess. The broken vertical line marks May 1649.

The immediate impression from the figure is easy to describe: for "equality", Levellers and non-Levellers are essentially constant in their similarity of understanding of this term (cosine distance ~ 0.3) until around 1657, when the similarity declines. But for "elections", there is a once-and-for-all increase in cosine similarity around 1647. That is, Levellers and non-Levellers use election terms in *more* similar contexts as time goes on. As statistical evidence, we conducted a Kendall Tau (monotonic) trend test for both series. As expected, the slope of the agreement for the "equality" terms is indistinguishable from zero, while for "election" terms, the slope is positive and statistically significant ($p = 0.068$).

6   Discussion

There has been a recent surge of interest in the Levellers, and for good reason: along with their broad appeal to political philosophers, as early theorists of democracy, the Levellers have ongoing appeal to liberals, libertarians and socialists alike. Yet many central interpretive questions remain. Here, we provided new data and new methods to help answer those questions. Instead of focussing on a few key works, we used the universe of extant Leveller pamphlets. This allowed us to situate the Levellers, and Lilburne particularly, in a much broader context than is usually possible. Specifically, we were able to study the works of this group relative to contemporaries who are *not* thought of Levellers, and to also look at how the Leveller arguments varied with events. In that sense, we made use of cross-section and time-series information.

Methodologically, we also innovated. We constructed word embeddings for our corpus that enabled us to talk about differences in 'meaning' across and between actors and times. This was challenging in part owing to the nature of the period and the sparsity of data—any language requires large amounts of data to produce useful embedding representations, yet we had only 260 pamphlets and Early Modern English is not necessarily easy to obtain in bulk. So we used parliamentary debates to build these vectors. From inspection, these appear valid and should be a useful tool for others.

As regards empirical findings, we had three primary ones. First, we confirmed that Lilburne was a linguistic innovator: he invented the "freeborn Englishman" construction—to the extent we can tell from our larger data—but our story is more subtle than other efforts. Specifically, we showed that his shift to this use was not "once and for all", and he tried various alternatives before the main one took off—both in his own work and that of others. Second, we showed that Leveller discussions of "equality" are primarily legal in nature, and differ from their contemporaries. Qualitatively, we showed that these "legal" rights are typically connected to notions of jury trial by "equalls". Lilburne stands apart from his Leveller colleagues, but the difference between the Levellers and their contemporaries is not driven only by "Freeborn John." Finally, the legal focus of the Levellers extends to their understanding of elections. That said, over time, the Levellers and their contemporaries become considerably more similar on the institutions of elections than they do on notions of equality.

Of course, our work above is preliminary. We have barely scratched the surface, even on the questions we have attempted to answer. For example, future efforts should check the sensitivity of our findings to different periodization and the inclusion or exclusion of different authors. Nonetheless, we now have the resources and an approach which can move these debates forward.

References

Anderson, Elizabeth. 2017. Private government. In *Private Government*, ed. Stephen Macedo. Princeton University Press.

Baker, Philip. 2013. *The Nature of the English Revolution Revisited: Essays in Honour of John Morrill*. Woodbridge: Boydell and Brewer chapter The Franchise Debate Revisited: The Levellers and the Army, pp. 103–22.

Bejan, Teresa M. 2022. "What Was the Point of Equality?" *American Journal of Political Science* 66(3):604–616. Publisher: John Wiley & Sons, Ltd.

Blaydes, Lisa, Justin Grimmer and Alison McQueen. 2018. "Mirrors for Princes and Sultans: Advice on the Art of Governance in the Medieval Christian and Islamic Worlds." *The Journal of Politics* 80(4):1150–1167.

Caliskan, Aylin, Joanna J Bryson and Arvind Narayanan. 2017. "Semantics derived automatically from language corpora contain human-like biases." *Science* 356(6334):183–186.

Camacho-Collados, Jose and Mohammad Taher Pilehvar. 2018. "From word to sense embeddings: A survey on vector representations of meaning." *Journal of Artificial Intelligence Research* 63:743–788.

De Krey, Gary Stuart. 2017. *Following the Levellers. Volume One, Political and Religious Radicals in the English Civil War and Revolution, 1645-1649*. London : Palgrave Macmillan UK.

Firth, John R. 1957. "A synopsis of linguistic theory, 1930-1955." *Studies in linguistic analysis* .

Foxley, Rachel. 2004. "John Lilburne and the Citizenship of 'Free-Born Englishmen'." *The Historical Journal* 47(4):849–874. Publisher: Cambridge University Press.

Foxley, Rachel, Peter Lake, Anthony Milton, Jason Peacey and Alexandra Gajda. 2016. *The Levellers : Radical political thought in the English Revolution*. Oxford: Oxford : Manchester University Press.

Frank, Joseph. 1955. The Levellers. In *The Levellers*. Harvard University Press.

Hart, DM and R. Kenyon. 2018. "Tracts on Liberty by the Levellers and their Critics (1638-1660), Cilt: 5,(Ed: DM Hart & R. Kenyon), Indianapolis: Liberty Fund.".
URL: *https://oll.libertyfund.org/title/hart-tracts-on-liberty-by-the-levellers-and-their-critics-1638-1660-7-vols*

Hill, Christopher. 1958. *Puritanism and Revolution*. Secker and Warburg.

Khodak, Mikhail, Nikunj Saunshi, Yingyu Liang, Tengyu Ma, Brandon Stewart and Sanjeev Arora. 2018. "A la carte embedding: Cheap but effective induction of semantic feature vectors." *arXiv preprint arXiv:1805.05388* .

Macpherson, C. B. 1962. *The political theory of possessive individualism: Hobbes to Locke*. Oxford: Clarendon Press.

Ober, Josiah. 2012. "Democracy's Dignity." *American Political Science Review* 106(4):827–846. Edition: 2012/09/13 Publisher: Cambridge University Press.

Pennington, Jeffrey, Richard Socher and Christopher D Manning. 2014. Glove: Global vectors for word representation. In *Proceedings of the 2014 conference on empirical methods in natural language processing (EMNLP)*. pp. 1532–1543.

Pocock, John. 1957. *The Ancient Constitution and the Feudal Law*. Cambridge University Press.

Rheault, Ludovic and Christopher Cochrane. 2020. "Word embeddings for the analysis of ideological placement in parliamentary corpora." *Political Analysis* 28(1):112–133.

Rodman, Emma. 2020. "A timely intervention: Tracking the changing meanings of political concepts with word vectors." *Political Analysis* 28(1):87–111.

Rodon, Toni and Tom Paskhalis. 2021. "Parliament Strikes Back: Agenda-setting and Power Voids in Early Representative Assemblies." *Working Paper* .
URL: *https://doi.org/10.31219/osf.io/qgu9c*

Rodriguez, Pedro L and Arthur Spirling. 2022. "Word embeddings: What works, what doesn't, and how to tell the difference for applied research." *The Journal of Politics* 84(1):101–115.

Rodriguez, Pedro L, Arthur Spirling and Brandon M Stewart. 2021*a*. "Embedding Regression: Models for Context-Specific Description and Inference.".

Rodriguez, Pedro L., Arthur Spirling and Brandon Stewart. 2021*b*. *conText: 'a la Carte' on Text (ConText) Embedding Regression*. R package version 1.0.0.

Schiffer, Stephen. 1972. *Meaning*. Oxford: Oxford University Press.

Seaberg, Robert B. 1981. "The Norman Conquest and the common law: the Levellers and the argument from continuity." *The Historical Journal* 24(4):791–806.

Skinner, Quentin. 1965. "I. History and Ideology in the English Revolution1." *The Historical Journal* 8(2):151–178.

Waldron, Jeremy. 2016. *Political Political Theory: Essays on Institutions*. Harvard University Press.