# The Origins of American Zoning

Alexander W. Bartik, Arpit Gupta, Daniel Milo, and Desmond O'Shaughnessy*

March 27, 2026

**Abstract**

We construct the first large-scale longitudinal dataset of American zoning ordinances, spanning nearly a thousand municipalities across one hundred years. We combine historical codes retrieved through direct municipal outreach, newspaper archives, and modern ordinances to document three central facts about the evolution of zoning. First, zoning has shifted away from strict Euclidean use segregation toward flexible and mixed-use provisions, with prohibitions on multifamily housing, accessory dwelling units, and mixed use declining substantially. Second, regulatory attention has migrated toward environmental, aesthetic, and growth-management objectives implemented through procedural requirements such as site plan review and design boards. Third, zoning codes have grown dramatically more complex: lengthening from roughly 20 to 200 pages, accompanied by more zoning districts and greater dispersion in bulk regulations across zones. These facts suggest that the impact of housing regulation is less due to persistence of early zoning rules than the accretion of novel regulatory goals layered into new procedural instruments and complex structures.

JEL-Classification: R52, R31, H73, D78, C55, C81

Keywords: housing regulation, zoning codes, housing supply, state and local government, large language models

*Bartik: Department of Economics, University of Illinois at Urbana-Champaign, abartik@illinois.edu. Gupta: New York University, Stern School of Business, ag5808@stern.nyu.edu. Milo: New York University, Stern School of Business, dm4766@stern.nyu.edu. O'Shaughnessy: New York University, djo9220@nyu.edu. We graciously acknowledge funding from Emergent Ventures, UIUC Campus Research Board, and the Chao-Hon Chen Institute for Global Real Estate Finance at NYU Stern which facilitated this research. We thank Omid Aliahmadi, Dev Gadde, Sophia Hong, Daniel Park, Saul Richardson, Rohan Varma, Shyam Naryanan, and Austin Kirwin for excellent research assistance.

# 1 Introduction

American zoning emerged in the early twentieth century as municipalities sought to bring order to rapidly industrializing cities, but its origins remain contested among legal scholars, urban historians, and economists. While the legal framework of zoning dates to the 1926 Supreme Court decision in *Village of Euclid v. Ambler Realty Co.* and early zoning codes, such as the 1916 New York City Ordinance, the economic and political drivers of housing regulations present in zoning codes are a matter of ongoing debate. To understand these forces, some scholars emphasize the role of zoning as a progressive force to address genuine market failures, for instance, separating noxious industrial uses from residential urban neighborhoods and coordinating urban development in the absence of Coasian bargaining (Fischel, 2015). Others highlight its role in advancing exclusionary impulses, noting how even early zoning codes restricted multifamily housing, protected single-family neighborhoods, and reinforced racial and class segregation (Rothstein, 2017; Cui, 2024). Understanding the origins and continued development of zoning regulations has implications for housing affordability, urban form, and the access to economic opportunity across American cities today (Glaeser and Gyourko, 2018).

A central challenge in tracing the persistent impact of American zoning is the lack of data on historical zoning ordinances themselves, which have served as the principal legal objects outlining local housing regulations. The fragmentation of American municipalities and the complexity of parsing long textual documents which encode housing regulations has long been an impediment to the scholarly analysis of zoning codes. To circumvent these challenges, some scholars have focused on in-depth analysis of zoning documents in specific areas (Glaeser and Ward, 2009). Others have focused on understanding cross-sectional patterns of housing regulations through survey-based approaches (Gyourko et al., 2008). These challenges are further magnified when analyzing historical zoning documents, which present additional challenges for retrieval and analysis. Typical historical zoning codes are stored in archives, with limited digitization and collection efforts conducted on this corpus at a nationwide level. To date, only a few papers have engaged in systematic longitudinal analysis of zoning regulations, typically in specific areas (for instance, focusing on Cook County as in Gallagher et al. (2025)). This presents an important

knowledge gap in understanding the nature of housing regulations over time, particularly at inception.

This paper directly addresses this gap on historical housing regulatory changes through a systematic collection and analysis of zoning records from a large cross-section of American municipalities over time using novel LLM methods. Our approach deploys a scalable generative AI approach to the analysis of regulatory documents, generative regulatory measurement as pioneered in Bartik et al. (2025), to a novel corpus of historical regulatory texts we assemble for this work.

Our central contribution is the creation of a dataset of initial housing regulations at the municipality level, spanning 976 cities across one hundred years to the present day, matched with modern zoning codes for 81% of the sample.[1] This provides the most comprehensive historical picture of American zoning from inception to the modern day to date. We do so by combining two novel data sources. The first is collecting zoning ordinances printed in local newspapers. The second source is through a comprehensive email campaign through which we reach out directly to municipalities in order to request and retrieve historical municipal ordinances. We combine these initial zoning codes with data on contemporary zoning codes for the same set of municipalities when available, generating a unique dataset that allows us to understand how zoning has evolved since its inception across a wide array of municipalities.

Using this dataset, we highlight three central facts about the evolution of American zoning.

**Fact 1: From use segregation to flexible and mixed uses.** Early zoning codes were principally organized around the Euclidean functions of separating land uses which achieved the ultimate objectives of preventing negative externalities, and promoting public health. We measure this by assessing the intent of zoning ordinances, based on an LLM categorization of document clauses, and find that early zoning codes were principally organized around the traditional "police power" using the constitutional authority of state and local governments to regulate private property for public health, safety, morals, and general welfare.

We complement our analysis of ordinance regulation intent with an analysis of regulation

[1] We plan to make the dataset publicly available.

type, which we separately classify using ordinance text.[2] Early ordinances were substantially concerned with land use and bulk regulations, which served the function of ensuring the strict segregation of land uses across residential, commercial, and industrial activities; as well as limiting the density and scope of residential development.

Over the course of the twentieth century, this emphasis has decreased substantially. Modern codes increasingly permit flexible and mixed uses across zones, including planned unit developments, mixed-use overlay districts, and accessory dwelling units. The restrictiveness of some bulk regulations, particularly on height limitations for districts allowing multifamily housing units, has declined. The share of text dedicated to these traditional Euclidean motives and implementations has decreased substantially over time.

Collectively, these findings indicate that housing regulations have actually liberalized along the use-segregation dimension over time. This has happened across municipalities regardless of the date they enacted their original zoning code, indicating that the persistence of these zoning codes is lower than commonly thought. It is also a surprising result in light of a common presupposition in the housing literature that housing regulations have become more binding for prices in the post-1970 period (e.g.. Glaeser et al., 2005). This shift is possibly the result of deindustrialization and the declining salience of nuisance externalities which lessened the need for traditional Euclidean separation of use functions; the potential benefits of agglomerative returns to mixing in service economies; and the potential shift of exclusionary functions to other regulatory instruments, as we focus on below.

**Fact 2: The rise of procedural and process-based regulation to achieve novel environmental, aesthetic, and fiscal management objectives.** As use-based restrictions have relaxed, regulatory attention has shifted toward environmental protection, aesthetic design control, and growth management—objectives that are accomplished in large part through process requirements. For example, site plan reviews, special permit procedures, environmental impact assessments, and design review boards have proliferated across the country. While municipalities have become

[2]In our framework, the type of regulation relates its physical legal content, whereas the intent of a regulation relates to what it seeks to accomplish. For example, a regulation containing a side setback to ensure access to light and fresh air would be classified by type as a bulk regulation and would be classified by intent as public health and safety.

more flexible on what can be built, they have simultaneously made it harder to navigate the process of building.

Our textual analysis shows that clauses classified under regulatory intents classified as environmental, fiscal, infrastructure coordination, and aesthetic grew substantially in both absolute number and in relative terms beginning in the 1950s, when the length of new regulatory ordinances began to grow substantially. These novel regulatory objectives largely account for the extensive margin expansion in the length of zoning ordinances.

**Fact 3: Growing complexity and dispersion.** Alongside these compositional shifts, zoning codes have become dramatically more complex. The median code has grown from roughly 25 pages in the earliest adoption cohorts (1900–1950) to over 75 pages for codes adopted from 1950 to 1980, and have grown drastically in size for codes since 1990, nearing 200 pages. This rise in length potentially reflects technological shifts in the production of documents as well as improvements in municipal state capacity that resulted in longer, more complex documents.

This expanded document length is reflected in regulatory complexity along a number of different dimensions. The number of distinct zoning districts within a single municipality has expanded markedly between historical and modern codes. Bulk regulations—minimum lot sizes, setback requirements, and height limits—have become more dispersed across districts, reflecting a more granular and heterogeneous regulatory landscape. Importantly, this shift in many regulations simultaneously reflects more binding regulations in the most restrictive zones, as well as less binding regulations in the least restrictive districts. The implication of these findings is that regulations have grown more targeted and precise in their application.

We emphasize a few caveats about our approach. First, and most notably, our data are subject to potential selection bias as a consequence of our email and newspaper-based retrieval approach. We measure the selection and bias of our sample in Section 2. Second, our study does not seek to analyze the causal determinants or subsequent impacts of regulatory shifts, which we leave for future work. Our focus in this paper is the descriptive analysis of housing regulations themselves, which are of independent interest as an object of study. And third, the regulations

we study are the municipal level based on regulatory texts themselves. In practice, what also matters substantially are the interpretation of these regulations, the available local state capacity to enforce them, the willingness of local planners to change the rules in response to local pressure; as well as the specific geographies to which bulk regulations apply (i.e., the actual zones themselves in accompanying zoning maps which we have also collected). We also leave for future work the important task of disentangling these different components to establish the degree to which printed zoning codes were binding.

Our work relates to three strands of literature. Most relevant is prior literature on zoning regulations (Fischel, 2002, 2015). For instance, Shertzer et al. (2018); Zhou et al. (2008); McDonald and McMillen (1998) analyze historical zoning regulation in the Chicago area. A common finding in this literature is that zoning codes, once adopted, tend to be remarkably persistent. Ellickson (2020) documents the "zoning straightjacket," that land use designations in established neighborhoods of detached houses are persistent, while Shertzer et al. (2022) emphasizes the slow-moving nature of land ordinances, and traces the impact of early ordinances on contemporary urban geography. While the housing stock itself may be persistent, reflecting the durable nature of structures built under earlier rules, we emphasize that the ordinances themselves appear to be considerably more dynamic. The textual content of zoning ordinances has changed substantially over time, with codes shifting away from rigid use segregation towards flexible mixed-use provisions, and regulatory attention moving towards environmental, aesthetic, fiscal, and later affordability requirements. Remarkably, we see relatively little evidence of differences in modern ordinances by the year of the initial zoning ordinance, despite the fact that initial zoning ordinances changed substantially over time.

Also relevant is a considerable literature on modern housing regulations (Huang and Tang, 2012; Gyourko et al., 2021), which has sought to understand the role of regulations in affecting housing markets by influencing both supply and demand (Mayer and Somerville, 2000; Gyourko and Molloy, 2015; D'Amico et al., 2023; Hsieh and Moretti, 2019; Duranton and Puga, 2019; Ganong and Shoag, 2017; Saiz, 2010). Our paper is principally focused on the question of the origins of modern housing regulations.

We also connect to a literature examining regulatory texts in general. Different approaches

to analyzing regulatory text include those of Al-Ubaydli and McLaughlin (2017); Singla (2023); Calomiris et al. (2020); Kalmenovitz et al. (2025). We differ by focusing on an aspect of local municipal regulation for which we are able to observe longitudinally, which enables us to measure aspects of persistence over time as well as cross-sectional features of regulatory adoption across the United States.

Finally, we also relate to a broader literature on the development of American cities (Boustan et al., 2013; Baum-Snow, 2007; Bleakley and Lin, 2012). A key question in this literature has been the role of local governance in connecting local economic fundamentals in shaping urban spatial outcomes. While much of this literature has focused on infrastructure investments, fiscal policy, and jurisdictional fragmentation as determinants of urban growth and sorting, less attention has been paid to the regulatory texts that directly govern the physical form of development.

# 2 Data

A core contribution of this paper is the construction of a unique dataset on the evolution of American municipal zoning codes since their inception. In this section, we start by providing key background on the history of American zoning. We then detail our two-pronged approach to collecting these historical data, which uses a combination of a national direct outreach campaign to municipalities and parsing of zoning ordinances posted in historical newspapers. Finally, we outline how we pair these historical zoning codes with the contemporary zoning code for each our sample municipalities.

## 2.1 Background on American Zoning

American zoning originated in the early twentieth century as municipalities sought legal tools to manage the spatial organization of rapidly growing cities. The 1916 New York City Zoning Resolution is typically cited as the first comprehensive zoning ordinance in the United States, though several cities had adopted partial land-use regulations earlier. The constitutional basis for comprehensive zoning was established by the Supreme Court in *Village of Euclid v. Ambler*

*Realty Co.* (1926), which upheld the authority of municipalities to enact zoning regulations as a legitimate exercise of police power. Following *Euclid*, states enacted enabling statutes that delegated zoning authority to local governments, and adoption spread rapidly: by 1930, zoning ordinances covered a majority of the urban population (Knauss, 1929), and they continued to spread in subsequent decades.

## 2.2 Direct Outreach to Municipalities

We conduct the first national outreach effort to collect historical zoning ordinances. Many municipalities still maintain archival copies of their original zoning ordinances (Gallagher et al., 2025). However, no comprehensive enumeration of historical zoning ordinances exists at either a state or national level, and the vast majority of local governments do not publicize catalogs of their own archival holdings. As a result, it is unclear how many and which municipalities maintain zoning archives. We show that direct municipal outreach, when paired with generative AI tools, makes possible the effective retrieval of archival municipal regulations.

We develop a municipal outreach algorithm detailed in Appendix Section A.1 and Figure A1, with the objective of retrieving the first comprehensive zoning ordinance and its accompanied map for every local government. We start by scraping emails for officials from each local government and ranking contacts by relevance.[3] We then send an initial outreach email to the highest priority contact (see Appendix Figure A2 for an example) and await a response. If a contact does not respond, we follow up or escalate to the next priority contact. Municipal responses fall into three categories: 1) a reason records are unavailable or costly to retrieve,[4] 2) a request for more information, which often includes filling out a records request form, and 3) a potentially relevant document. When a document is received, we evaluate whether it constitutes the municipality's first comprehensive zoning ordinance; if so, we mark the municipality as complete, and if not, we use clues within the document—such as references to earlier ordinances by adoption

[3] We rank contacts with an LLM that has the objective to prioritize first town clerks, then zoning officials, and finally other staff.

[4] Records may not be available if they were lost or destroyed. Other towns present costs for document retrieval. This includes towns that require in-state residency for records requests (i.e., Tennessee) and municipalities that charge money for record retrieval, which we skip.

date—to redirect our request.

We build a scalable human-AI hybrid system to execute our outreach algorithm on over the universe of local governments. First, we automate routine parts of the pipeline like sending follow up emails to towns that haven't replied within a week. We then rely on generative AI for both simple classification tasks and first passes at more complex tasks. For example, we fully rely on AI to collect municipal contacts, classify the town for inbound emails, and parse ordinance PDFs.[5] For more complex tasks like drafting emails, filling records request forms, and classifying inbound documents, we use AI to draft responses for human review. To involve human assistance we create a ticket system where human research assistants provide judgment/approval when needed. We begin by focusing on local governments from the 2022 Census of Governments with a population of at least 10,000. Thus far we have reached out to about a third of the governments in this sample and expect to scale to the rest as well as to smaller population localities.

Our resulting sample exhibits limited observable bias relative to the national universe of municipalities. Table 1 compares our email outreach sample to the broader universe of municipalities from which it is drawn. On most observable characteristics, our sample exhibits limited bias relative to the national comparison group. The most notable differences are regional: our sample slightly over-represents municipalities in the Northeast and West and under-represents those in the Midwest and South.[6] Our sample also skews toward somewhat larger and more metropolitan municipalities, consistent with the fact that larger municipalities are more likely to maintain organized archives and dedicated staff to respond to records requests.

## 2.3 Publication of Zoning Ordinances in Newspapers

Historical zoning ordinances were also frequently printed in local newspapers, which constitutes a second source of data for our project. This was the result of two main legal drivers. First,

[5] For parsing we use a Vision Language Model, specifically Chandra 2 OCR.

[6] In the Midwest, many townships and municipalities have overlapping jurisdictional boundaries, and in a number of cases only one of the two entities is responsible for zoning. Some of the lower Midwest coverage in our sample reflects the fact that contacted township governments informed us that they do not administer zoning. Residency requirements also limit our ability to retrieve samples from some regions, particularly in the South.

some state enabling statutes that provide authority for local zoning came with requirements for public hearing, advance notices, and legal notices typically in a "newspaper of general circulation." Second, many states require ordinances of a general nature to be published as a condition of being effective. The definition of qualifying newspaper varied, but was frequently defined by states as a function of frequency, circulation, length of publication, etc. for the purpose of official status. These requirements were frequently interpreted by municipal counsel to entail publishing the entire zoning ordinance, particularly for the first such code or for a comprehensive rewrite (see Appendix Figure A4 for examples).

We draw on the website `newspaperarchive.com` as a source for the historical newspaper corpus from which we focus on complete zoning ordinance text. We start by identifying newspaper issues that may contain relevant information by doing keyword searches for issues that contained the three terms "zoning", "ordinance", and "dwelling." We then select pages of the issues containing one of these key words and used Vision Language Models (VLMs) to generate machine readable text from the pages. We classify the materials based on whether they included a complete ordinance, a partial or abridged ordinance, an amendment, a notice, or non-ordinance material discussing zoning (or several of these categories) and focus on the issues containing full ordinances.[7] We restrict the set of ordinances derived from newspapers that we include in our sample to full ordinances that were classified as having less than 5 percent of the ordinance missing. Appendix Section A.2 and Figure A3 provide an overview of the entire process of constructing this set of historical ordinances from newspapers.

## 2.4 Modern Ordinances

For municipalities where we obtain historical zoning codes through our email and newspaper collection strategies, we also collect the current ordinances. The addition of these modern ordinances allows us to understand how zoning codes have changed from the inception of zoning to the modern day for each of our sample municipalities. For each municipality in our historical sample, we search for a match in the set of modern codes collected by Bartik et al. (2025) along

[7] In future work, we plan on incorporating material other than full ordinances into our analysis.

with newly collected codes from eCode.com. The resulting match rate is 81.4%.

## 2.5 Municipal Sample Characteristics

We now investigate the characteristics of the unique sample of historical municipal zoning ordinances that we constructed following the procedure described above. Our sample includes 765 municipalities from 49 states who responded to our direct outreach efforts by providing their initial zoning code and 232 municipalities for whom we found a full historical zoning ordinance in a newspaper archive. Combined, this yields a sample of 976 unique municipalities for which we have their first zoning ordinance.[8]

Figure 1 shows maps of the number of municipalities included in our sample from each of these two sources. The top map in this figure shows that the email ordinance collection approach finds ordinances from municipalities spread throughout the country and different kinds of counties. The bottom map shows that ordinances found in newspapers were somewhat more regionally concentrated, with a particularly large number of these ordinances concentrated in the upper Midwest and fewer coming from municipalities in the South. These two maps show that when combined, these two sources generate a sample of historical zoning ordinance with remarkably good coverage of all states and county types within the United States.

We then investigate the characteristics of sample municipalities in more detail in Table 1. For both the email outreach and newspaper archives sub-samples, the table reports municipal coverage and incorporation year (Panel A), region (Panel B) metropolitan and urban status (Panel C), and modern economic characteristics (Panel D) both overall and for the relevant comparison group. The comparison group for the email sample contains all municipalities with populations greater than 10,000 and less than 200,000, consistent with our email outreach criteria. The newspaper comparison group contains all municipalities in the Census of Governments. Our samples diverge somewhat from the relevant universe of municipalities on a few dimensions. Both samples exhibit moderate selection toward larger, more metropolitan, and higher-income municipalities compared to the relevant full universe. Relative to the relevant universe of munic-

[8]There were 21 overlapping municipalities for whom we found ordinances through both the email and newspaper archive approach.

ipalities, we somewhat over-represent the Northeast and West and under represent the Midwest and South. However, overall, both samples are fairly similar to the relevant comparison population on most dimensions.

Table 2 then investigates the timing of adoption and length of initial zoning ordinances by urban status, region, and time. The mean adoption year in our full sample is 1960, with urban municipalities adopting roughly six years earlier than suburban ones, and the Northeast and Midwest leading the South and West, likely reflecting in part later incorporation for municipalities in these regions. The sample includes a large number of municipalities whose initial codes were authored in every time period, allowing us to explore how codes varied with the cohort of adoption. As we will discuss in more detail below, zoning codes have lengthened considerably over time. The median initial codes were 31 pages, but this varied substantially by era: codes adopted before 1940 averaged less than 20 pages, while those adopted after 1980 averaged over 88 pages.

# 3 Empirical Strategy

We take two broad approaches to measuring the content of zoning regulations over time. First, we divide the zoning text into clauses and classify both the broad type and intent of each clause. These broad classifications capture how the scope, goals, and emphasis of zoning regulations have changed over time. Second, we measure a large set of specific features of the zoning regulation using the set of regulation measures from Bartik et al. (2025). These regulations, which include both categorical variables such as whether townhouses are permitted and numerical regulations such as the maximum height limit, were chosen to represent the most important municipal zoning rules.

We now lay out each of these approaches in more detail.

## 3.1 Clause Classification

Zoning codes are written to achieve objectives that change over time. As political priorities shift, regulatory technologies develop, and broader intellectual currents change, the substance of what zoning regulations regulate and the purposes they serve can change substantially. Capturing these compositional shifts requires us to distinguish between two dimensions of regulatory text.

The first is the type of regulation: what the regulation governs. A specific regulatory clause might regulate land use (specifying which activities are permitted within a zone), impose bulk and dimensional standards (height limits, setbacks, or minimum lot sizes), establish procedural requirements (mandate site plan review or public hearings), or address other domains such as parking, environmental protection, or signage.

The second dimension is intent: what a regulation is designed to accomplish. This can be a challenging inference because regulations may achieve multiple purposes. For instance, a bulk regulation might be intended to manage stormwater runoff in one municipality, preserve rural character in another, or exclude low-income households in a third. Conversely, a single regulatory objective, such as growth management, may be pursued through multiple regulatory types, such as bulk regulations limiting density or procedural rules requiring phased developmental approvals. Because both type and intent may vary independently, it is important to track both dimensions to establish how scope and goals have changed since inception. We do so using an LLM-based classification which categorizes individual code clauses along both dimensions.

Our classification begins by first dividing up each code into clauses. We do this using a standard set of legal clause markers (such as “SECTION" or "I.", see a full list of these clauses in Appendix Section A.3). Second, we prompt the LLM to classify each clause into one of a predetermined set of eleven types and ten zoning intents. To create this selection of types and intent, we formed an initial set using *a priori* knowledge from the literature and then fine-tuned our selection through a careful reading of 30 randomly selected codes. The types and intents used within this classification are described below.

**Type classification.** Each clause is assigned to one of the following categories: *land use restrictions* (primary and accessory use regulations), *bulk restrictions* (setbacks, lot sizes, height limits,

floor area ratios), *procedural* (governing body powers, variance process, appeals, special permits, site plan review, enforcement), *nonconforming use* (grandfathering provisions), *definitions*, *parking*, *environmental protection* (habitat, stormwater, erosion, environmental overlay districts), *historical preservation*, *flexible zoning* (planned unit developments, open space residential design, mixed-use overlays), *inclusionary zoning* (affordable housing requirements, density bonuses), and *miscellaneous* (fences, signs, etc.). Specific details on prompting are given in Appendix Section A.4.

**Intent classification.** Each clause is also classified by regulatory intent: *exclusion* (restrictions on density, housing types, or populations), *externality control* (separation of incompatible uses), *public health and safety*, *infrastructure coordination*, *environmental protection*, *growth management*, *value capture* (exactions, dedications), *fiscal zoning* (attracting commercial tax base), and *aesthetic/design control*. Specific details on prompting are given in Appendix Section A.5.

With these categories in hand, we submit groupings of ten consecutive clauses to `gemini-3-flash` and prompt it to categorize each clause, using the surrounding clauses as context. We do this separately for the type and intent categorization. Table 3 shows examples of different combinations of these type and intent categorizations on clauses from our historical sample.

Two patterns stand out from the evolution of initial zoning codes over time. First, the length of initial codes has exploded, rising from an average of under twenty pages for municipalities whose first code was released in the 1920s to over 100 pages for those adopted in the 1990s. Second, the composition of codes has shifted markedly: new intents such as environmental protection have emerged and grown in prominence, while traditional objectives such as public health and safety have declined as a share of the total code. Figure 2 documents both patterns, showing the distribution of clause topics (Panel A) and intents (Panel B) for initial zoning codes by adoption cohort. The left graph in each panel shows the absolute number of pages devoted to each category, illustrating the growth in overall code length, while the right graph shows the share of the total code allocated to each category.

### 3.2 Specific Regulatory Extraction

We measure specific regulations following the generative regulatory measurement approach developed in Bartik et al. (2025). That study demonstrated that LLMs achieve near-human accuracy in classifying zoning regulations, with 96% accuracy on binary questions and a 0.87 correlation on continuous questions. Methodologically, the approach entails (1) identifying the relevant sections of an ordinance for a question, (2) preparing guidelines, instructions, and steps for each question, and (3) determining the optimal LLM hyperparameters for regulatory analysis. We carry over the set of regulatory questions developed in that study, originally drawing on classifications from (Glaeser and Ward, 2009) and key regulations identified in (Turner et al., 2014), along with the background information and prompting strategies that were iteratively refined. We also extend these to additional questions relevant to historical codes. These regulations include use restrictions determining permitted land uses, bulk regulations governing building density (such as minimum lot sizes, setbacks, and height limits), process rules for approving new development, and regulatory mandates such as affordable housing requirements and environmental provisions.

Unlike Bartik et al. (2025), which uses Retrieval-Augmented Generation (RAG) to identify the relevant sections of each ordinance, we instead feed the entire historical ordinance to the LLM. This is made possible by the smaller sample size and page length of historical ordinances as well as new discounts by several LLM providers on repeated long-context queries.[9]

## 4 Results

Our data on initial and contemporary zoning code allows us to investigate three dimensions of the development of American zoning codes. First, we can measure how initial zoning codes changed over time. This helps us understand how zoning practices of municipalities newly implementing a zoning ordinance have changed over time.

Second, we investigate how zoning ordinances have changed between the initial code and the contemporary code within municipality. This approach surmounts a key limitation of the first

[9] Specifically, `gemini-3-flash` offers a 90% discount on repeated queries on the same context.

analysis: the composition of municipalities initiating zoning codes changes over time, making it difficult to separate cohort effects from genuine regulatory shifts.

Third, we combine these two types of analysis and investigate how the within municipality change in zoning ordinances between the initial and contemporary codes varies with the year of the initial code, which we refer to as the "ordinance cohort." Exploring the within municipality change in zoning practices by cohort helps us understand how sticky initial zoning practices have been.

We summarize these findings in our first fact:

## 4.1 Fact 1: From use segregation to flexible and mixed uses.

The earliest American zoning codes were organized around a central principle inherited from *Euclid*: the strict separation of incompatible land uses. Residential, commercial, and industrial activities were assigned to distinct zones, with the stated objectives of preventing nuisance externalities and protecting public health. Over the subsequent century, this Euclidean foundation has eroded substantially. Modern codes increasingly permit flexible and mixed uses, allow a wider range of housing types, and devote a shrinking share of their text to traditional use segregation. We document this transformation using both the clause-level classification and the specific regulatory measures described above.

In Figure!3, we show how the topics (left sub-figure) and intents (right sub-figure) of clauses in zoning codes have changed between municipalities' first zoning code and their current zoning code, averaged across all adoption cohorts. The share of the code devoted to each topic or intent is shown as a turquoise circle for the initial code and as a dark blue triangle for the modern code. Reflecting the broad shifts we saw in Figure 2 above, we see a marked shift in the share of zoning codes devoted to different topics and intents. The first zoning codes in municipalities were dominated by land-use regulation and bulk regulations. The intent of these clauses was primarily to ensure public health and safety, had the effect of excluding some types of residents or uses, and was intended to control externalities. These initial topics and intents have become much less central to zoning ordinances over time, with the share of the code devoted to bulk

regulations declining from over 15 to 10 percent and on land-use restrictions declining from 18 to 15 percent. The change in emphasis in terms of clause intent has been even more marked, with the share of the code devoted to externality control or public health and safety declining from around 45 percent to around 20 percent.

One possibility is that the declining share of code text dedicated to traditional use regulation could simply reflect dilution with additional topics, rather than genuine liberalization. To test this possibility, Figure 4 shows the difference between initial and modern codes across the twenty one specific regulations that we measure. We find evidence of genuine use liberalization: on almost every dimension, modern zoning codes allow more flexible uses, including a falling share that prohibits mixed use (38 to 18 percent), mobile homes (55 to 27 percent), attached single family (50 to 18), Accessory Dwelling Units (ADUs, 83 to 43), and multi-family (55 to 17 percent). The one measured use regulation for which there was no substantial increase in flexibility was building multi-family by right, which was roughly unchanged.

In Figure 6, we explore how these changes in zoning between the first code and modern codes have varied by the year in which the municipality created its first zoning code, which we refer to as its "zoning cohort." This investigation provides a more granular picture of how zoning practices have evolved, and helps us understand the extent to which a municipalities' original zoning practices are sticky over time. In this figure, we see that these use restrictions were particularly stringent for municipalities whose initial zoning codes were written earlier. For example, for municipalities whose initial zoning codes were written between 1920 and 1940, over 80 percent did not allow multifamily development through a flexible zoning process, over 90 percent prohibited mobile homes, over 80 percent prohibited ADUs, and over 60 percent prohibited attached single family. Generally, the initial zoning codes were more flexible the later the original code was written. Strikingly, the modern codes for most of these use restrictions are similar across municipalities from different zoning cohorts (with the notable exception of mobile home restrictions). This pattern suggests that zoning may not be extremely sticky and municipalities have adapted zoning readily to reflect changing goals or economic conditions.

We summarize these findings with our first fact:

**Fact 1.** *Zoning codes have shifted away from segregation of uses/use restrictions aimed at preventing negative externalities, public health, and safety problems, and instead are now increasingly allowing more flexible and mixed uses across zones.*

## 4.2 Fact 2: The rise of procedural regulation in pursuit of novel regulatory objectives.

In Fact 1, we detailed the types of regulation that were declining in emphasis and stringency. In this subsection, we investigate which regulations have risen in importance. Returning to Figure 3, we see that while clause intents related to externality control and public health and safety were declining in importance, others were rising in importance or being implemented for the first time. In particular, clauses whose intents related to environmental protection, growth management, aesthetic and design control, and infrastructure coordination have risen from forming less than 15 percent of the content of zoning codes to almost 40 percent. Strikingly, if we use the more granular temporal information on the intents of first codes in Figure 2, we see that before 1950, zoning codes essentially did not contain clauses whose intent related to environmental protection or value capture. Other zoning intents, such as aesthetic design and control always existed, but were relatively trivial parts of initial zoning codes for early zoning cohorts.

How have these new zoning intents been implemented? Figure 7 shows the probability that a clause is of a certain type conditional on the clauses's intent. We see that clauses related to these rising zoning intents are disproportionately discussed in clauses related to process rules. For example, 79 percent of clauses whose intent is growth management are from procedural sections of the zoning code. Appendix Table A1 confirms these results via a multinomial logit of clause types on clause intents.

This development is also evident in specific regulatory measures. Returning to Figure 4, we see that almost all of the regulations that have increased in stringency the most are procedural regulations. In most cases, process rules have gone from being quite rare—for example, almost

no municipalities required site plan review in the original zoning code—to being present in the majority of municipalities. These changes are quite large across these process regulations, with requiring impact statements, site plan review requirements, public hearing requirements, and prohibiting use variances all rising in frequency by over 20 percentage points. Similarly, specific regulations related to the new regulatory goals discussed above have also become much more common. For example, environmental or historical district overlays were almost never present in municipalities' original zoning codes but are now present in a large share of municipalities.

Turning to Figures 8 and 9, which show how zoning codes have evolved for municipalities whose first zoning codes were released in different cohorts, we see that strict procedural rules and these novel zoning intents were much less common among earlier zoning codes. Indeed, they were almost non-existent in most cases. However, municipalities from all original code cohorts have converged on a relatively similar rate of regulations imposing stringent procedural rules and implementing new regulatory intents related to aesthetics and the environment.

These patterns give rise to our second stylized fact:

**Fact 2.** *Zoning codes have shifted into environmental, aesthetics, and growth management goals that are often accomplished through process restrictions and rules.*

### 4.3 Fact 3: Growing complexity and dispersion.

One of the most striking patterns of Figure 2 is the dramatic rise in zoning code length. In particular, after 1970 the average length of codes has grown substantially corresponding with the rise of the word processor and the acceleration of land use court cases documented in Ganong and Shoag (2017). This pattern is even more striking when comparing municipalities' initial code length to their modern code length. In Figure 4, we see that for the average municipality, between their first zoning code and now the number of pages has risen by almost four standard deviations and the number of districts has risen by two standard deviations, translating into a rise of roughly 150 pages and 12.9 districts. Figure 10 depicts this change in length of codes and

number of districts for cohorts. Average code lengths rise from around 20 pages to 50 pages between the 1920–1940 and 1960–1980 cohorts, rise to 120 pages in the 1980–2000 cohort, and increase to over 200 pages in the modern era. This pattern is also reflected in the average number of zones, which starts at around 6 for early cohorts and rises to around 20 for all cohorts.

Table 4, specification (1) shows that nearly a third of this increase in code length is accounted for by the rise in the number of zoning districts. Each zoning district is associated with a code that is more than seven pages longer. In specification (2), we include changes in regulatory indices that measure use stringency, process oversight, and additional goals based on the regulations shown in Figures 6, 8, and 9, respectively. The oversight and additional goals indices increase as municipalities add more process restrictions and regulatory goals and are also associated with substantial increases in page counts. The use stringency index, which increases with use stringency, is negatively correlated with an increase in pages, which likely reflects that including more and flexible uses requires additional regulatory content. These facts suggest these longer zoning codes reflect the need for more length to describe regulations related to additional goals, more process oversight, and more flexible uses in a larger number of districts.

Municipalities have used these larger number of districts to enact finer and more complicated zoning regulations. In Figure 4, we see that across municipalities the maximum bulk regulation across districts rose in every bulk regulation category, while the minimum was unchanged or shrank. Thus the range of bulk regulations has risen substantially within municipalities between their initial regulations and today. These longer codes have allowed municipalities to develop much finer zoning regulations targeted to particular places within the municipality.

Figure 11 shows how these changes from initial to modern zoning code have varied by zoning cohort. We see that these patterns are especially stark for early initial zoning codes. The initial codes in these municipalities had relatively narrow gaps between the minimum and maximum bulk regulations across zones, suggesting that in its early phases many ordinances were relatively simple. These gaps have widened significantly over time, suggesting that zoning has become much more targeted to particular locations.

We summarize this pattern with our third stylized fact:

**Fact 3.** *Zoning codes have become much more complex: they're longer, have more zones, and the rules across zones have become more dispersed.*

# 5 Conclusion

Our paper presents the first large-scale longitudinal analysis of American zoning ordinance texts. By combining direct municipal outreach and newspaper-published historical codes with applying LLM-based textual analysis, we construct a novel panel dataset that tracks the content and structure of zoning regulations from their early twentieth-century origins to the present day. This creates the first comprehensive and systematically analyzed record of historical housing regulations across the United States.

Our analysis reveals three central facts. First, American zoning has shifted away from strict use segregation toward more flexible and mixed-use provisions, reflecting a decline in the original Euclidean origins of zoning in the separation of incompatible uses. Second, regulatory attention has migrated toward environmental, aesthetic, and fiscal growth-management objectives implemented principally through procedural requirements. Third, codes have grown dramatically in length and complexity over time, reflecting technological and state capacity shifts in document production, which enable more granular regulatory treatment of bulk regulations across areas.

These results challenge some common prevailing views that zoning regulations are unusually persistent in stringency or have become uniformly more restrictive over time. On one important margin, the traditional Euclidean separation of uses that characterized the earliest codes, regulation has actually substantially liberalized. Additionally, the regulations in the least restrictive zoning districts have loosened along with increases in complexity and the number of zones. However, this liberalization has coincided with the proliferation of procedural requirements which raise the time, cost, and uncertainty of development, even as some of the underlying use permissions have expanded. We link these developments to shifts in regulatory attention away

from traditional externality and public health oriented motives for regulation, towards a broader set of regulatory intentions which municipalities sought to achieve through zoning regulations.

We leave for future work the ultimate drivers of these shifts in regulatory attention and complexity. One view is that civil rights legislation and judicial standards may have made explicit use-based exclusion more challenging to sustain legally, driving regulatory energy towards procedural instruments that achieve similar exclusionary objectives through less legally vulnerable mechanisms. Alternatively, the ultimate drivers for these trends may be somewhat distinct. The rise of the environmental movement and fiscal challenges for cities in the wake of deindustrialization and suburban flight may have generated novel regulatory impulses achieved through the zoning code.

# References


**Al-Ubaydli, Omar and Patrick A McLaughlin**, "RegData: A numerical database on industry-specific regulations for all United States industries and federal regulations, 1997–2012," *Regulation & Governance*, 2017, *11* (1), 109–123.

**Bartik, Alexander, Arpit Gupta, and Daniel Milo**, "The costs of housing regulation: Evidence from generative regulatory measurement," Technical Report, Working Paper 2025.

**Baum-Snow, Nathaniel**, "Did highways cause suburbanization?," *The quarterly journal of economics*, 2007, *122* (2), 775–805.

**Bleakley, Hoyt and Jeffrey Lin**, "Portage and path dependence," *The quarterly journal of economics*, 2012, *127* (2), 587–644.

**Boustan, Leah Platt, Devin Michelle Bunten, and Owen Hearey**, "Urbanization in the United States, 1800-2000," Technical Report, National Bureau of Economic Research 2013.

**Calomiris, Charles W, Harry Mamaysky, and Ruoke Yang**, "Measuring the cost of regulation: A text-based approach," Technical Report, National Bureau of Economic Research 2020.

**Cui, Tianfang**, "Did Race Fence Off The American City? The Great Migration and the Evolution of Exclusionary Zoning," Technical Report, Working Paper 2024.

**D'Amico, Leonardo, Edward L Glaeser, Joseph Gyourko, William Kerr, and Giacomo AM Ponzetto**, "Why Has Construction Productivity Stagnated? The Role of Land-Use Regulation," Technical Report, Working Paper 2023.

**Duranton, Gilles and Diego Puga**, "Urban growth and its aggregate implications," Technical Report, National Bureau of Economic Research 2019.

**Ellickson, Robert C**, "The zoning straitjacket: The freezing of American neighborhoods of single-family houses," *Ind. LJ*, 2020, *96*, 395.

**Fischel, William**, *Zoning Rules!: The Economics of Land Use Regulation*, Lincoln Institute of Land Policy, 2015.

**Fischel, William A**, *The homevoter hypothesis: How home values influence local government taxation, school finance, and land-use policies*, Harvard University Press Cambridge, MA, 2002.

**Gallagher, Ryan, Allison Shertzer, and Tate Twinam**, "Zoning and the American Suburb," November 2025. Preliminary Slides.

**Ganong, Peter and Daniel Shoag**, "Why has regional income convergence in the US declined?," *Journal of Urban Economics*, 2017, *102*, 76–90.

**Glaeser, Edward and Joseph Gyourko**, "The economic implications of housing supply," *Journal of Economic Perspectives*, 2018, *32* (1), 3–30.

**Glaeser, Edward L and Bryce A Ward**, "The causes and consequences of land use regulation: Evidence from Greater Boston," *Journal of Urban Economics*, 2009, *65* (3), 265–278.

_ , **Joseph Gyourko, and Raven E Saks**, "Why have housing prices gone up?," *American Economic Review*, 2005, *95* (2), 329–333.

**Gyourko, Joseph, Albert Saiz, and Anita Summers**, "A new measure of the local regulatory environment for housing markets: The Wharton Residential Land Use Regulatory Index," *Urban Studies*, 2008, *45* (3), 693–729.

_ **and Raven Molloy**, "Regulation and housing supply," in "Handbook of regional and urban economics," Vol. 5, Elsevier, 2015, pp. 1289–1337.

_ , **Jonathan S Hartley, and Jacob Krimmel**, "The local residential land use regulatory environment across US housing markets: Evidence from a new Wharton index," *Journal of Urban Economics*, 2021, *124*, 103337.

**Hsieh, Chang-Tai and Enrico Moretti**, "Housing constraints and spatial misallocation," *American Economic Journal: Macroeconomics*, 2019, *11* (2), 1–39.

**Huang, Haifang and Yao Tang**, "Residential land use regulation and the US housing price cycle between 2000 and 2009," *Journal of Urban Economics*, 2012, *71* (1), 93–99.

**Kalmenovitz, Joseph, Michelle Lowry, and Ekaterina Volkova**, "Regulatory fragmentation," *The Journal of Finance*, 2025, *80* (2), 1081–1126.

**Knauss, Norman L.**, "Zoning Progress in the United States: Zoned Municipalities in the United States," Technical Report, Department of Commerce, Division of Building and Housing, Washington April 1929.

**Mayer, Christopher J and C Tsuriel Somerville**, "Land use regulation and new construction," *Regional Science and Urban Economics*, 2000, *30* (6), 639–662.

**McDonald, John F and Daniel P McMillen**, "Land values, land use, and the first Chicago zoning ordinance," *The Journal of Real Estate Finance and Economics*, 1998, *16* (2), 135–150.

**Rothstein, Richard**, *The Color of Law: A Forgotten History of how our Government Segregated America*, Liveright Publishing, 2017.

**Saiz, Albert**, "The geographic determinants of housing supply," *The Quarterly Journal of Economics*, 2010, *125* (3), 1253–1296.

**Shertzer, Allison, Tate Twinam, and Randall P Walsh**, "Zoning and the economic geography of cities," *Journal of Urban Economics*, 2018, *105*, 20–39.

_ , _ , **and** _ , "Zoning and segregation in urban economic history," *Regional Science and Urban Economics*, 2022, *94*, 103652.

**Singla, Shikhar**, "Regulatory costs and market power," Technical Report 2023.

**Turner, Matthew A, Andrew Haughwout, and Wilbert Van Der Klaauw**, "Land use regulation and welfare," *Econometrica*, 2014, *82* (4), 1341–1403.

**Zhou, Jian, Daniel P McMillen, and John F McDonald**, "Land values and the 1957 comprehensive amendment to the Chicago zoning ordinance," *Urban Studies*, 2008, *45* (8), 1647–1661.

# Tables

Table 1: Sample Comparison Table

| | Email Outreach | | Newspaper Archives | |
|---|---|---|---|---|
| | Census Comparison | Our Sample | Census Comparison | Our Sample |
| **Panel A. Coverage** | | | | |
| Municipalities (N) | 6299 | 765 | 38736 | 232 |
| States represented | 50 | 49 | 51 | 35 |
| Year incorporated, median | 1895 | 1892 | 1900 | 1876 |
| **Panel B. Regional shares** | | | | |
| Northeast share (%) | 19.6 | 28.8 | 16.5 | 20.7 |
| Midwest share (%) | 33.8 | 25.5 | 55.8 | 54.3 |
| South share (%) | 31.9 | 23.4 | 21.0 | 10.3 |
| West share (%) | 14.8 | 22.4 | 6.7 | 14.7 |
| **Panel C. Metropolitan and urban status** | | | | |
| In MSA share (%) | 70.4 | 83.7 | 43.0 | 62.1 |
| Urban share (%) | 30.1 | 27.4 | 8.3 | 13.4 |
| Suburban share (%) | 62.9 | 64.1 | 15.6 | 47.0 |
| Rural share (%) | 7.0 | 8.5 | 76.1 | 39.6 |
| **Panel D. Modern census characteristics (2022)** | | | | |
| Population, median | 22,086 | 19,022 | 1,042 | 7,168 |
| Median household income ($) | 78,305 | 83,221 | 66,397 | 65,963 |
| Owner-occupied share (%) | 66.7 | 67.6 | 77.4 | 70.3 |
| Poverty rate (%) | 11.5 | 10.7 | 12.1 | 12.2 |

*Notes:* This table depicts summary statistics comparing our full historical email and newspaper samples to comparable sets of municipalities from the 2022 Census of Governments. The comparison group for the email sample contains all municipalities with populations greater than 10,000 and less than 200,000, consistent with our email outreach criteria. The newspaper comparison group contains all municipalities in the Census of Governments sample. The region is defined using U.S. Census regions (Northeast, Midwest, South, West). The urban/suburban/rural distinction is based on the census variable % urban. We classify urban as >95%, suburban as >20% and <95%, and rural as <20% .

Table 2: Adoption of Initial Codes by Urban Status, Region, and Over Time

| | Full sample | Urban | Suburban | Rural |
|---|---|---|---|---|
| **Panel A. By urban status** | | | | |
| Jurisdictions (N) | 976 | 226 | 567 | 147 |
| Mean adoption year | 1959.6 | 1953.9 | 1959.6 | 1970.1 |
| Share adopted by 1940 | 17.2 | 26.1 | 17.3 | 2.7 |
| Share adopted by 1960 | 54.5 | 65.9 | 56.4 | 27.2 |
| Share adopted by 1980 | 87.4 | 91.6 | 85.2 | 87.1 |
| Number of Pages (median) | 30.9 | 37.3 | 32.9 | 3.0 |
| **Panel B. By region** | | | | |
| | Northeast | Midwest | South | West |
| Jurisdictions (N) | 263 | 312 | 200 | 201 |
| Mean adoption year | 1957.1 | 1956.0 | 1967.7 | 1960.6 |
| Share adopted by 1940 | 21.7 | 21.5 | 10.0 | 11.9 |
| Share adopted by 1960 | 58.9 | 57.7 | 46.5 | 51.7 |
| Share adopted by 1980 | 91.3 | 93.6 | 72.5 | 87.6 |
| Number of Pages (median) | 29.1 | 20.9 | 47.7 | 38.8 |
| **Panel C. Over time** | | | | |
| | 1900-1930 | 1930-1950 | 1950-1970 | 1970-present |
| Jurisdictions (N) | 81 | 221 | 392 | 282 |
| Number of Pages (median) | 19.8 | 21.2 | 30.7 | 88.9 |

*Notes:* This table contains summary stats related to our full historical sample of initial zoning codes. Region defined using U.S. Census regions (Northeast, Midwest, South, West). The urban/suburban/rural distinction is based on the census variable % urban. We classify urban as >95%, suburban as >20% and <95%, and rural as <20% .

Table 3: Clause Examples by Intent and Type

| City | Intent | Type | Clause Text |
|---|---|---|---|
| Midlothian, TX 1959 | externality control | land use restriction | (C) Commercial District. – (1) Any use permitted in any of the foregoing districts. – (2) Blacksmith or horseshoeing shops... (34) Manufacture of any kind not listed under this section, provided that such use is not noxious or offensive by reason of the emission of dust, smoke, gas, noise or vibration... |
| Shreveport, LA 1958 | externality control | miscellaneous | (2) A six- to eight-foot-high solid fence of cedar, redwood, ... or solid wall of masonry or concrete shall be located along the side or rear property line that is adjacent to a residential district. The screening shall be adequate to obstruct views, reduce noise, divert emissions, or lessen other potential land use conflicts with adjacent properties... |
| Providence, RI 1923 | public health and safety | bulk restriction | SIDE AND REAR YARDS. Except in an unrestricted area district, side yards of a minimum width of 3 feet are required for all buildings erected in dwelling house or apartment house districts... Light and air are essential for health and comfort and each building should contribute its fair share to the light, air and open spaces of the entire block. |
| Lansing, IL 1955 | public health and safety | procedural | 9. ANNUAL INSPECTION-COMPLAINT OF BUILDING IN DANGEROUS CONDITION-AUTHORITY OF COMMISSIONER: The Building Commissioner and the Fire Commissioner shall make inspections of public institutions, ... If inspection shows that the building is unsafe for the public, the Commissioner shall notify the owner... and said changes or additions shall be made within 30 days from notice... |
| Glendora, CA 1932 | exclusion | land use restriction | 1. No building, structure, improvement, or premises shall be erected, constructed, established, altered, or enlarged in Zone "R-1", which is designed, arranged or intended to be occupied or used for any purposes other than a single-family residence... |
| Kissimmee, FL 1957 | exclusion | bulk restriction | (d) Building Area: No one-story principal dwelling shall be erected in a Residence "A-2" district with less than seven hundred and fifty (750) square feet of first floor area. No one and one-half story principal dwellings shall be erected in a Residence "A-2" district with less than six hundred and seventy-five (675) square feet of first floor area ... |
| Endicott, NY 1968 | infrastructure coordination | parking | **Sec. 20-68. Parking space.** In R-1 Residence Districts: - (a) Every one-family dwelling shall provide offstreet parking space for not less than one motor vehicle. - (b) For places of assembly there shall be not less than one offstreet parking space for each eight (8) persons... |
| Oviedo, FL 1971 | infrastructure coordination | definitions | 6. Limited Access Highway or Freeway - is a highway which permits no access except at authorized and controlled points, the acquisition of rights-of-way for such highway usually including the acquisition of access rights thereto. Access may also be limited through methods other than acquisition of access right. |
| Saratoga, CA 1961 | aesthetic design control | miscellaneous | Section 10.1 PURPOSES In order to preserve the natural beauty of the city's site, to maintain the orderliness of the community's appearance..., the location, size, illumination and design of signs are regulated ... No sign, outdoor advertising structure or display shall be erected, ... |
| Pewaukee, WI 1963 | environmental protection | land use restriction | (3) PROHIBITED USES. Dumping, filling, drainage, or any other use that would disturb the natural flora, character of the water bottom, land structure or topography including motor and ice boating, except to and from their moorings. |
| Sammamish, WA 1999 | growth management | procedural | 20.24.197 Additional Examiner findings and recommendations - school capacities. Whenever the Examiner in the course of conducting hearings... of potential multi-family zoning, receives documentation that the public schools in the district where the development is proposed would not meet the standards set out in ISDC 21A..., the Examiner shall remand to the Department to require or recommend phasing ... or deny the proposal... |

*Notes:* This table contains a selection of clauses categorized by our clause categorization pipeline. All clauses are selected from our full historical sample.

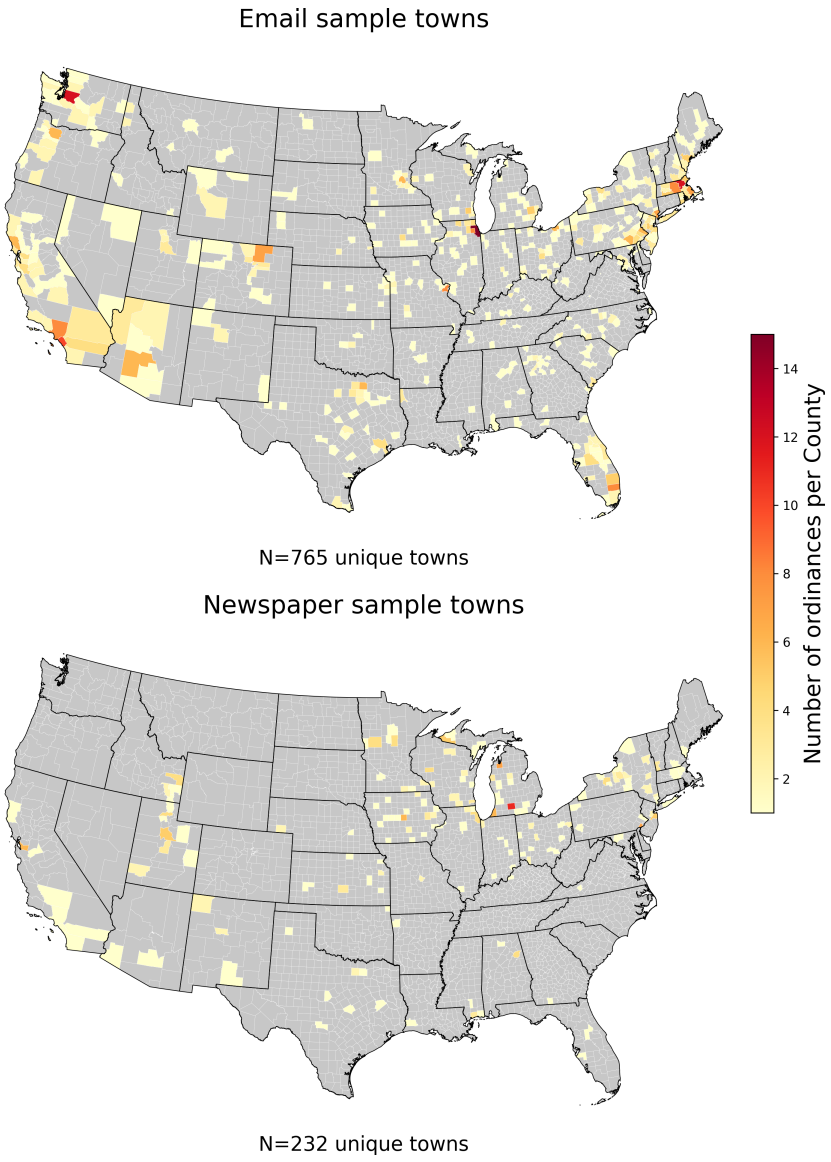
Table 4: Change in Page Count Regressed on District Change and Changes in Regulatory Indices with Cohort Fixed Effects

| | Zone count only | Zone count + indices |
|---|---|---|
| Change in number of districts | 7.64 | 5.93 |
| | (0.53) | (0.52) |
| Change in use stringency index | | 41.56 |
| | | (20.12) |
| Change in oversight index | | 93.48 |
| | | (18.16) |
| Change in additional goals index | | 114.85 |
| | | (21.70) |
| Cohort FEs | Yes | Yes |
| Constant | -132.96 | -270.22 |
| | (101.35) | (95.53) |
| Observations | 649 | 649 |
| $R^2$ | 0.356 | 0.443 |

*Notes:* This table shows the results of a regression of change in number of pages on change in number of zones (Column 1) and change in number of zones and regulatory indices (Column 2) from historical to matched modern codes. The use stringency index is composed of the five questions present in Figure 6, with a positive change in the index indicating that the municipality has gotten more restrictive when it comes to use. The procedural oversight index is composed of procedural questions from Figure 8, with higher values indicating stricter procedural regulations. The additional goals index is composed of the questions present in Figure 9, with higher values associated with additional regulatory goals. All indices are set on a scale from zero to one. Cohort decade fixed effects are included in both specifications. The sample includes the subset of historical and matched modern codes, with some slight attrition due to LLM uncertainty about number of zones in certain historical and modern codes.

# Figures

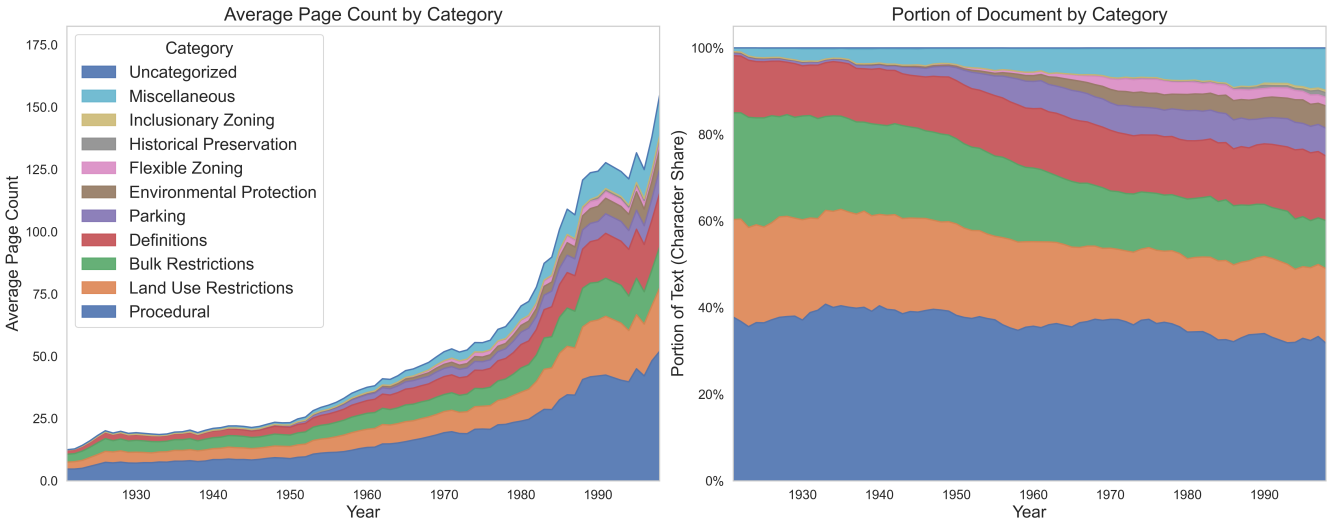
Figure 1: Historical Municipal Zoning Ordinance Coverage by County



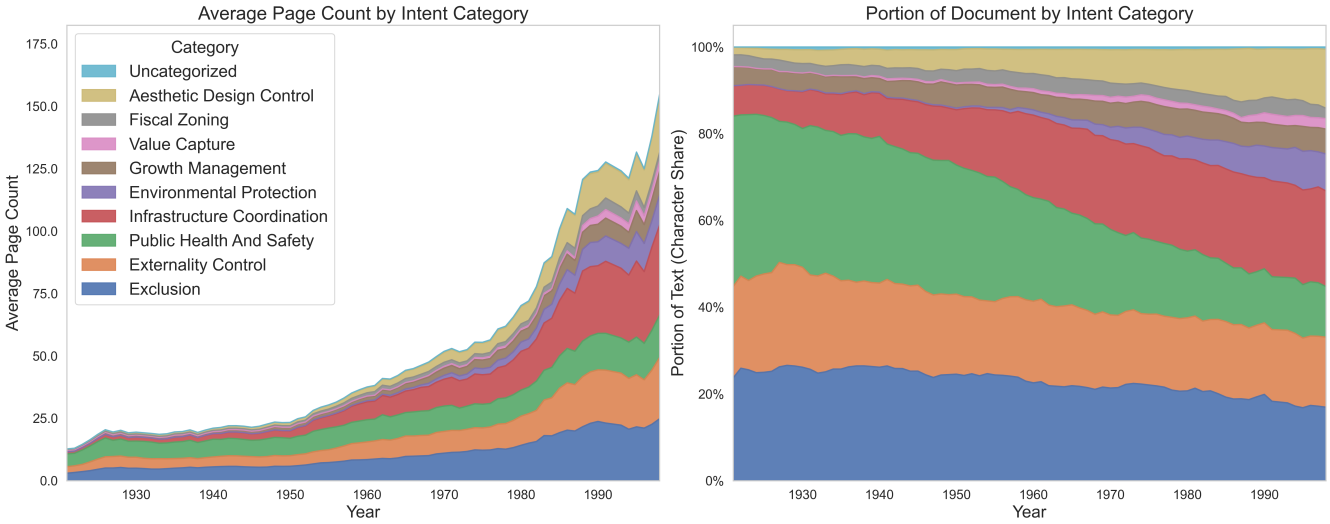

*Notes:* This figure is a county level map depicting the number of towns in each county for which we have an initial zoning code.

Figure 2: Topical and Intent Composition of Zoning Code Clauses Over Time

Panel A: Clause Category

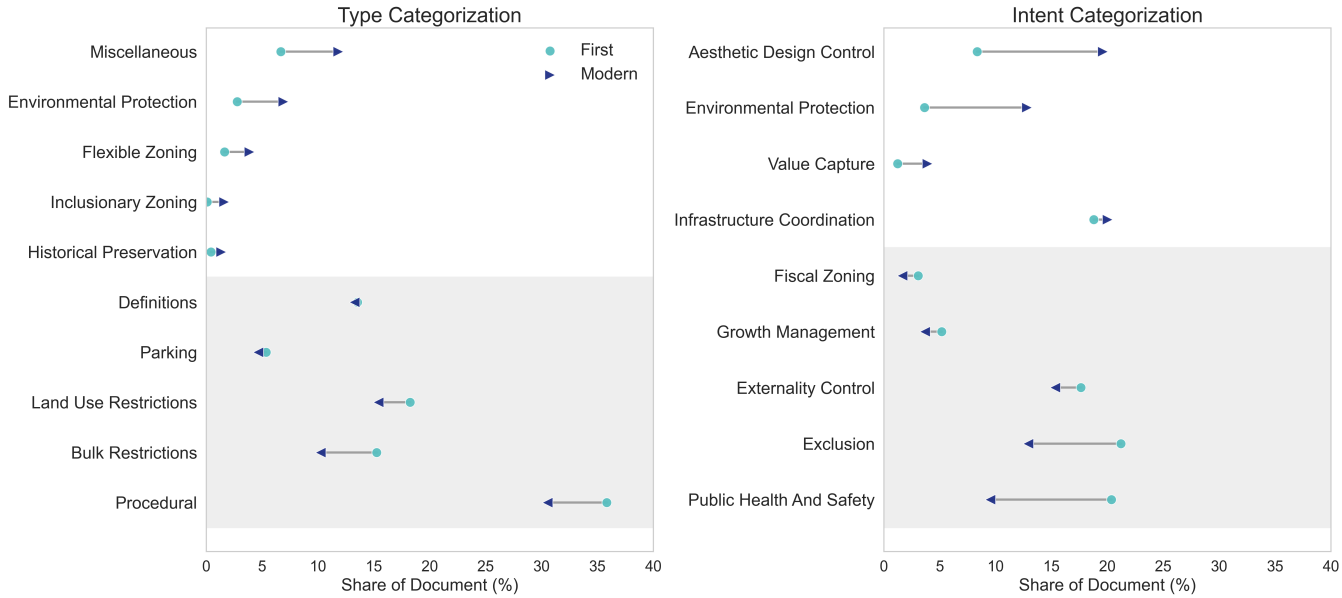



Panel B: Clause Intent



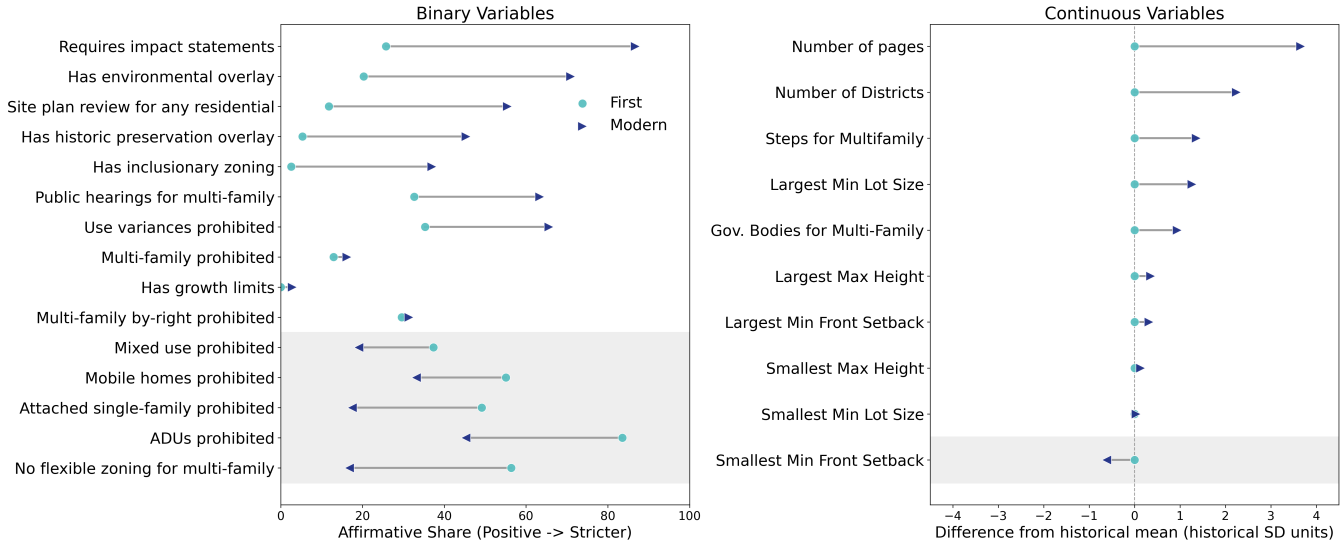

*Notes:* Each panel shows absolute (left) and relative (right) proportions. Documents are split into clauses using a regex approach with common heading and legal clause conventions, then classified by an LLM. Panel (a) assigns each clause to a topical category (e.g., land use restrictions, bulk restrictions, procedural, environmental protection, flexible zoning). Panel (b) assigns each clause to an intent category (e.g., exclusion, externality control, growth management, aesthetic/design control). Figures represent ten-year rolling averages. Page counts on the left plots reflect character counts divided by 2000 in order to ensure consistency across differently formatted documents. The sample includes all historical ordinances.

Figure 3: Clause categorizations: Change from First to Modern



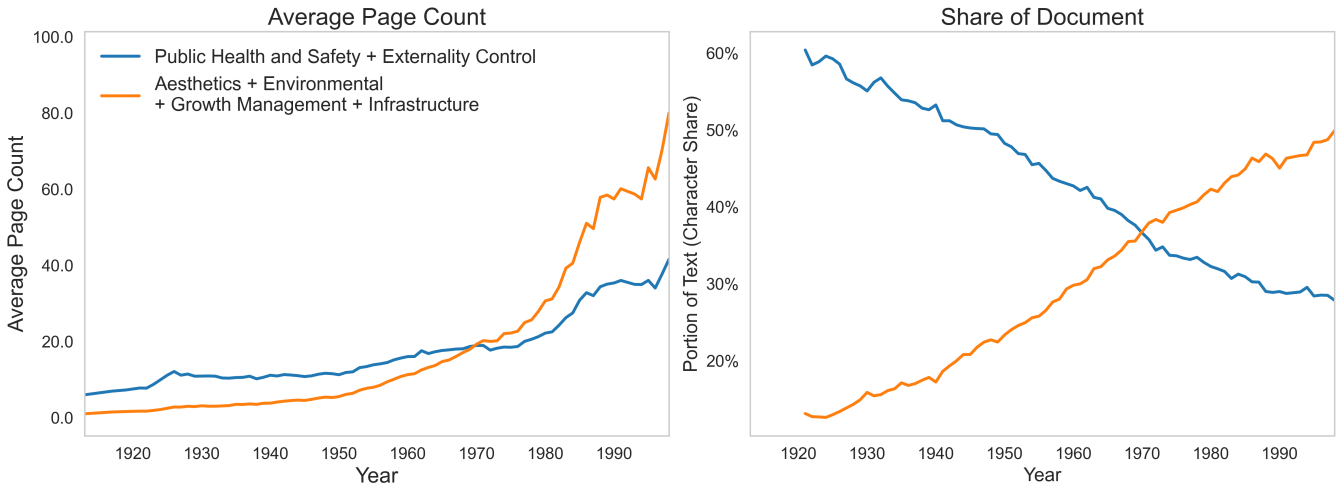

*Notes*: This figure depicts the change in document share of each type (left) and intent (right) category from the pool of all initial codes to their modern counterparts. The categories are sorted by change. The sample includes the subsample of historical codes for which we have a modern match.

Figure 4: Binary and continuous questions: change from first to modern



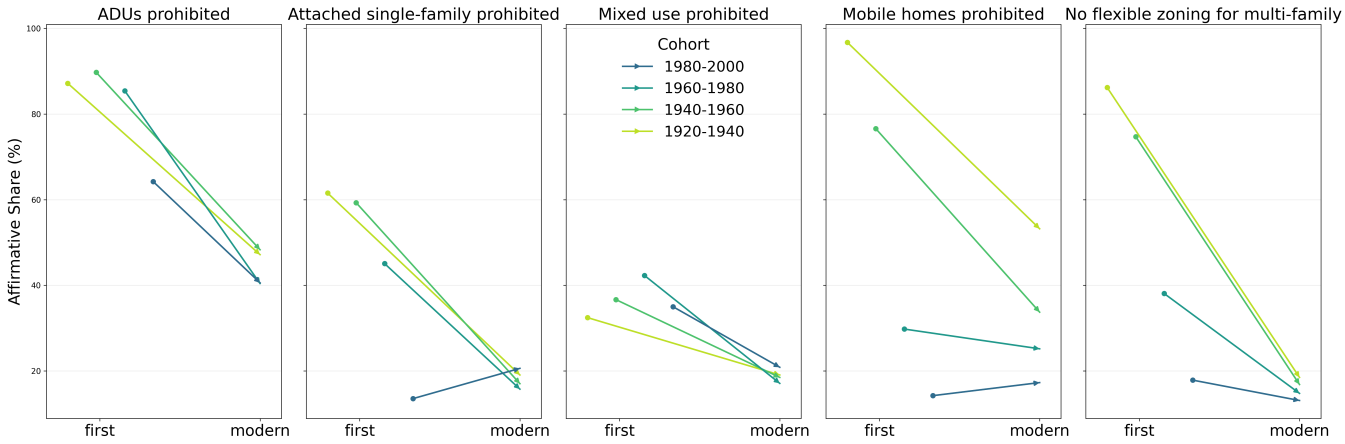

*Notes*: This figure depicts the change in binary questions (left) and continuous questions (right) from the pool of all initial codes to their modern counterparts. The left panel shows change in share of codes that are affirmative, with higher values associated with stricter zoning policies. The right panel depicts changes in continuous variables by comparing the average levels in the modern codes to the standard deviation of their historical counterparts. The categories are sorted by change.The sample includes the subsample of historical codes for which we have a modern match.

Figure 5: Public health, safety, and externality vs. other intent clauses over time



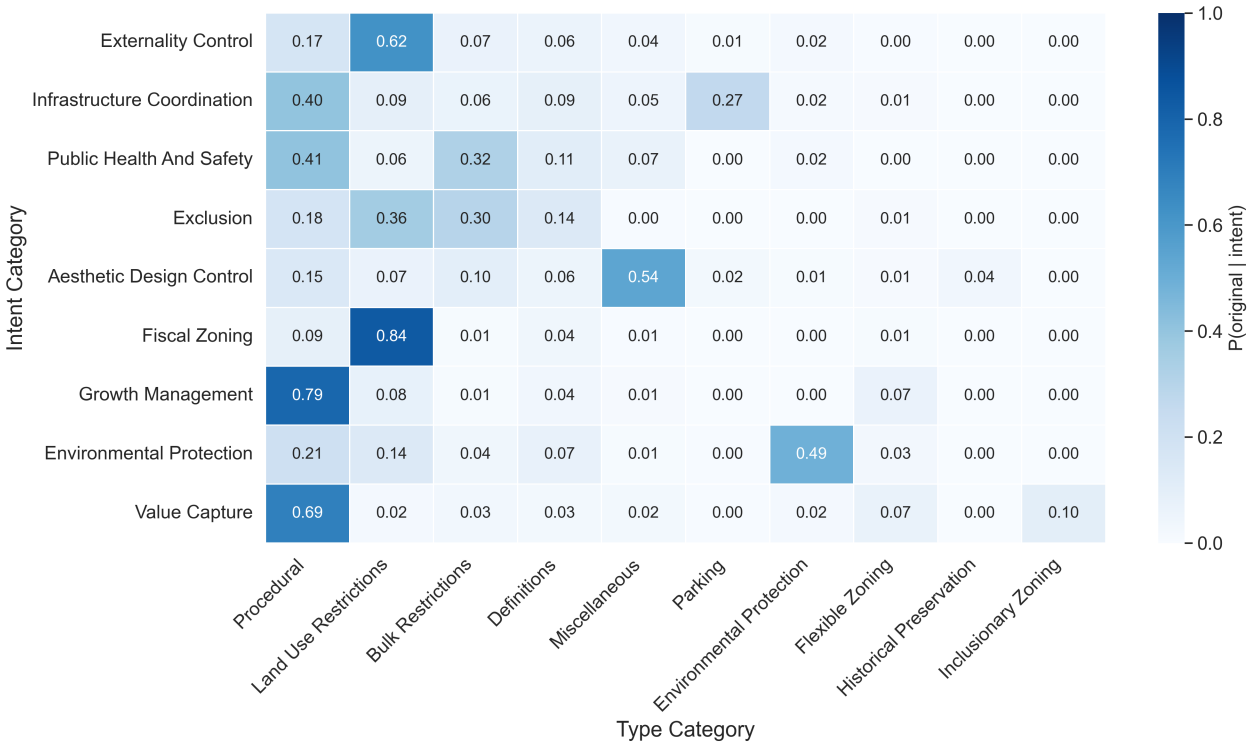

*Notes:* This figure shows the average page count (left) and document share (right) of different type categorization groupings. The sample includes all historical zoning codes.

Figure 6: Use regulations over time: historical vs matched modern



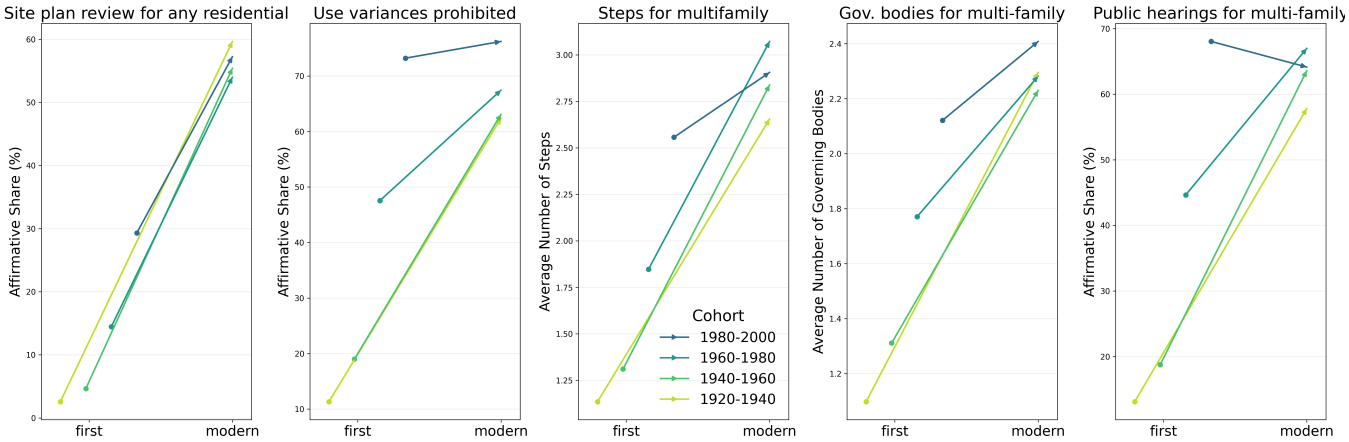

*Notes:* This figure depicts changes in selected use regulations from historical zoning codes to their modern counterpart by zoning-code inception cohort. The sample includes the subsample of historical codes for which we have a modern match.

Figure 7: P(Type Category || Intent Category)



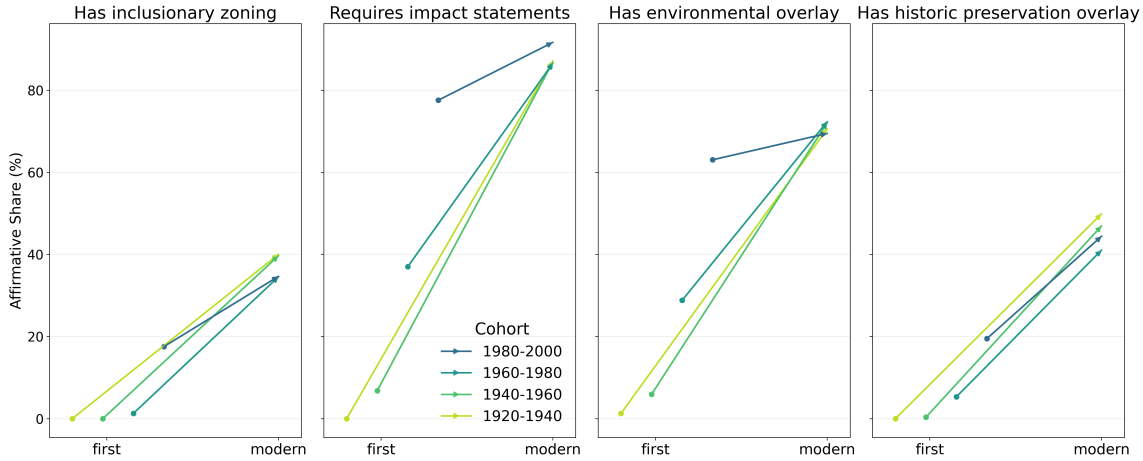

*Notes:* This figure shows the probability of a clause being categorized a certain type given its intent category. The sample includes all historical ordinances.

Figure 8: Procedural regulations over time: historical vs. matched modern



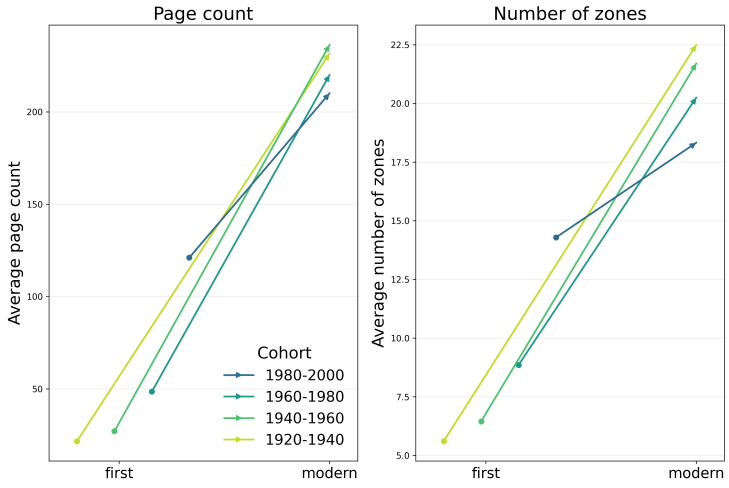

*Notes:* This figure depicts changes in selected procedural regulations from historical zoning codes to their modern counterpart by zoning-code inception cohort. The sample includes the subsample of historical codes for which we have a modern match.

Figure 9: Additional regulatory goals over time: historical vs. matched modern



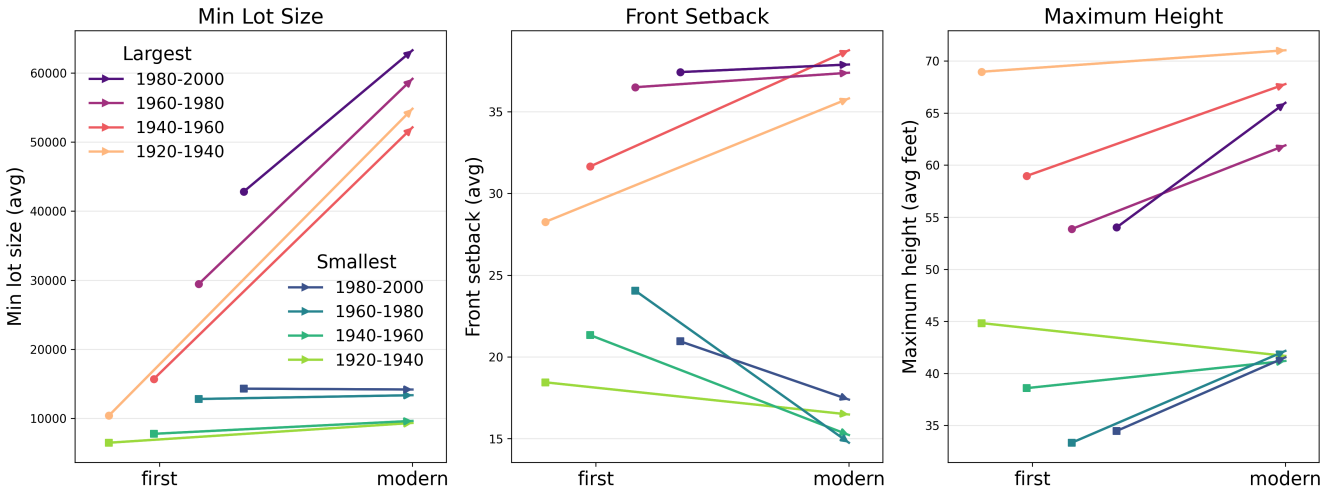

*Notes:* This figure depicts changes in regulations with additional goals from historical zoning codes to their modern counterpart by zoning-code inception cohort. The sample includes the subsample of historical codes for which we have a modern match.

Figure 10: Zone and Page Count: historical vs. matched modern




*Notes:* This figure depicts changes in number of zones and page count from historical zoning codes to their modern counterpart by zoning-code inception cohort. Page counts on the right plot reflect character counts divided by 2000 in order to ensure consistency across differently formatted documents.

Figure 11: Min lot size, front setback, and max height over time: historical and matched modern




*Notes:* This figure depicts the changes in the average highest and lowest front setback, minimum lot size, and max height over time from historical zoning codes to their modern counterpart by zoning-code inception cohort. The setback and lot size regulations depicted are for single family residential zones: zones that are primarily residential and allow single family but not multi-family developments. The max height is for any zone that allows residential.

# A Additional Details on Methodology

This appendix section contains additional details on our sample creation and empirical methodologies. In Sections A.1 and A.2, we outline the process of obtaining our email and newspaper samples in greater detail. In Section A.3, we provide details on the process used to extract legal clause chunks from zoning ordinances. In Sections A.4 and A.5, we show the prompts used to categorize extracted clauses by regulation type and intent, respectively.

## A.1 Email Outreach

Appendix Figure A1 details our email outreach pipeline. We begin by using web agents to find municipal contacts. Specifically, we query the Serper API for each town's contact page and then use an LLM to navigate and extract contact information from the resulting search results. A second LLM pass ranks these contacts, prioritizing town clerks followed by zoning and planning staff.

We then send a personalized initial email to the highest-ranked contact requesting the municipality's earliest zoning ordinance and accompanying zoning map (see Appendix Figure A2). If no response is received after one week, we send a follow-up email; if another week passes with no response, we escalate to the next highest-priority contact with a new initial outreach email.

Inbound replies are automatically classified by an LLM. When a municipality requests a formal records request, the system fills the required form—either a PDF attachment or a web-based FOIA portal—without manual intervention. When documents are received, each attachment is independently classified to determine its document type (i.e. full zoning ordinance, amendment, adoption shell, comprehensive plan, etc.) and date.

All cases requiring human judgment—such as composing non-routine replies, evaluating received documents, or deciding whether to pursue earlier records—are escalated to a research assistant via a web-based review portal. The portal displays the full correspondence history, the LLM's classification and suggested draft reply, and all received documents with their extracted metadata. The research assistant can then send a reply, redirect outreach to a different contact, mark the municipality as complete, and review document classifications.

## A.2 Newspaper Parsing

We draw on NewspaperArchive.com to assemble a historical newspaper corpus and identify issues likely to contain complete zoning ordinances. As shown in Appendix Figure A3, the pipeline begins with issue-level keyword retrieval: we first retain newspaper issues containing all three search terms "zoning", "ordinance", and "dwelling", and then collect the relevant page images from those candidate issues. We next apply OCR to the full selected pages[10] in order to recover page-level machine-readable text while preserving page structure for later reconstruction.

We then classify each issue using the page images together with the OCR text, distinguishing full ordinance publications from excerpts, amendments, notices, and non-ordinance zoning discussion, with the key requirement that a page contain clause-structured legal text rather than only summaries, captions, or hearing notices. For issues classified as likely full ordinances, the final reconstruction stage jointly uses the selected page images and OCR text to repair OCR errors, restore reading order across columns and page continuations, recover headings and numbering, and isolate only the operative ordinance text while excluding unrelated material. We keep in the sample only reconstructed newspaper ordinances classified as full ordinances with less than 5 percent of the text missing due to either OCR errors or newspaper abbreviations.

## A.3 Clause Chunking

In order to chunk clauses for classification, we use a regex based approach to identify legal language related to structural headings and clause markers. During the process of converting pdfs to markdown, the OCR tools identify many of the relevant headings and markers and categorizes them as such in the markdown output. However, because this process is imperfect, we fall back on regex to identify additional headers and clauses based on the following set of markers:

- Structural labels: TITLE, PART, CHAPTER, ARTICLE, SECTION, SEC, SEC., DIVISION
- Parenthesized numbers: (1), (2), ...

[10] Specifically we use Paddle OCR VL 1.5

- Parenthesized letters: (a), (b), . . .
- Parenthesized Roman numerals: (i), (ii), (iv), . . .
- Number-dot markers: 1., 2., . . .
- Letter-dot markers: a., b., . . .
- Dotted numeric markers: 1.1, 1.2.3, . . .
- Short all-caps headings: e.g. GENERAL PROVISIONS

## A.4 Type Categorization Prompting

We employ the following prompt for the LLM clause type categorization:

"You are a careful legal analyst. Categorize clauses using only the provided ordinance text.

You will categorize ordinance clauses from one window.

Return ONLY valid JSON with this schema: { "results": [ { "clause_id": 123, "category": "string", "confidence": 0.0 } ], "notes": "string" }

Rules:

- Include one result per `clause_id` provided.
- Do not omit clauses.
- Confidence must be a number between 0 and 1.

Categorization directions:

Pick one "category" for each clause based on its text content. Use the following categories:

1. `land_use_restrictions`
    - legislation that restricts how land can be used, such as primary uses, accessory uses, conditional uses, etc.
2. `bulk_restrictions`
    - legislation that restricts the physical form of development, such as height limits, floor area ratio limits, setback requirements, lot coverage limits, etc.
3. `procedural`

- description and powers of governing body, variance process, appeals process, special permit process, site plan review, certificates of occupancy, nonconforming use, enforcement and penalty, etc.

4. `definitions`
    - definitions of terms used in the ordinance.
5. `parking`
    - parking requirements, parking ratios, parking minimums, etc.
6. `environmental_protection`
    - habitat protection, drainage and stormwater, erosion, environmental overlay districts, environmental impact studies.
7. `historical_preservation`
    - protection of historical sites, historical preservation overlay districts, etc.
8. `flexible_zoning`
    - planned unit development, open space residential design, mixed use overlay districts, or other provisions that allow for flexibility in development.
9. `inclusionary_zoning`
    - requirements to build affordable units or density bonuses for affordable units.
10. `miscellaneous`
    - other regulations including aesthetics, fences, signs, courts, etc."

## A.5 Intent Categorization Prompting

We employ the following prompt for the LLM clause intent categorization:
"You are a careful legal analyst. Classify each clause's primary regulatory intent using only the provided ordinance text.
Return ONLY valid JSON with this schema:
{"results": ["clause_id": 123,"category": "string","confidence": 0.0}],"notes": "string"}

Rules: - Include one result per clause_id provided. - Do not omit clauses. - Confidence must be a number between 0 and 1.

Categorization directions: Pick one "category" for each clause based on its primary regulatory intent. Use only these categories:

1. `externality_control`
    - description: Separating incompatible land uses to prevent negative spillovers; protecting residential areas from industrial nuisances like noise, smoke, traffic, and odors; classic Pigouvian rationale for zoning
    - keywords: nuisance/noise/odor/dust references, separation buffers, industrial setbacks, smoke stack/boiler rules, hazardous storage, noxious, industrial, manufacturing, smoke, fumes, factory, noise, vibration, obnoxious, offensive, hazardous
2. `exclusion`
    - description: Restricting housing types, density, or building form in ways that raise housing costs and limit access; protecting property values and neighborhood character through supply constraints
    - keywords: single-family-only districts, multifamily bans or caps, minimum lot size threshold, family definition (related/unrelated caps; roomers/boarding restrictions), density limits, detached dwellings only, maximum units
3. `value_capture`
    - description: Extracting developer contributions to fund public goods or affordable housing; conditioning development approval on exactions, fees, or community benefits
    - keywords: impact fees, system development charges, dedication/parkland exactions, development agreements, inclusionary zoning, linkage fees, density bonuses, TIF
4. `growth_management`
    - description: Controlling the pace, timing, or location of development; limiting sprawl; preserving open space; managing infrastructure capacity; often framed around environmental or fiscal sustainability
    - keywords: moratorium, adequate facilities, phased development, growth boundary, urban limit, concurrency, buildout, carrying capacity
5. `fiscal_zoning`
    - description: Maximizing municipal tax base while minimizing service costs; favoring commercial/industrial over residential, or high-value housing over moderate-cost units
    - keywords: tax base, commercial district, retail, shopping center, industrial park, revenue, ratables

6. `public_health_and_safety`

    - description: Ensuring adequate light, air, ventilation for building occupants; fire prevention; sanitation focused on occupant welfare rather than neighbor externalities
    - keywords: light and air, ventilation, fire district, setback for fire, building code, sanitary, health, overcrowding, tenement

7. `aesthetic_design_control`

    - description: Regulating visual appearance of buildings, signage, and landscapes; architectural standards; historic preservation
    - keywords: sign, billboard, architectural review, design standards, historic district, facade, landscaping, scenic, visual blight

8. `infrastructure_coordination`

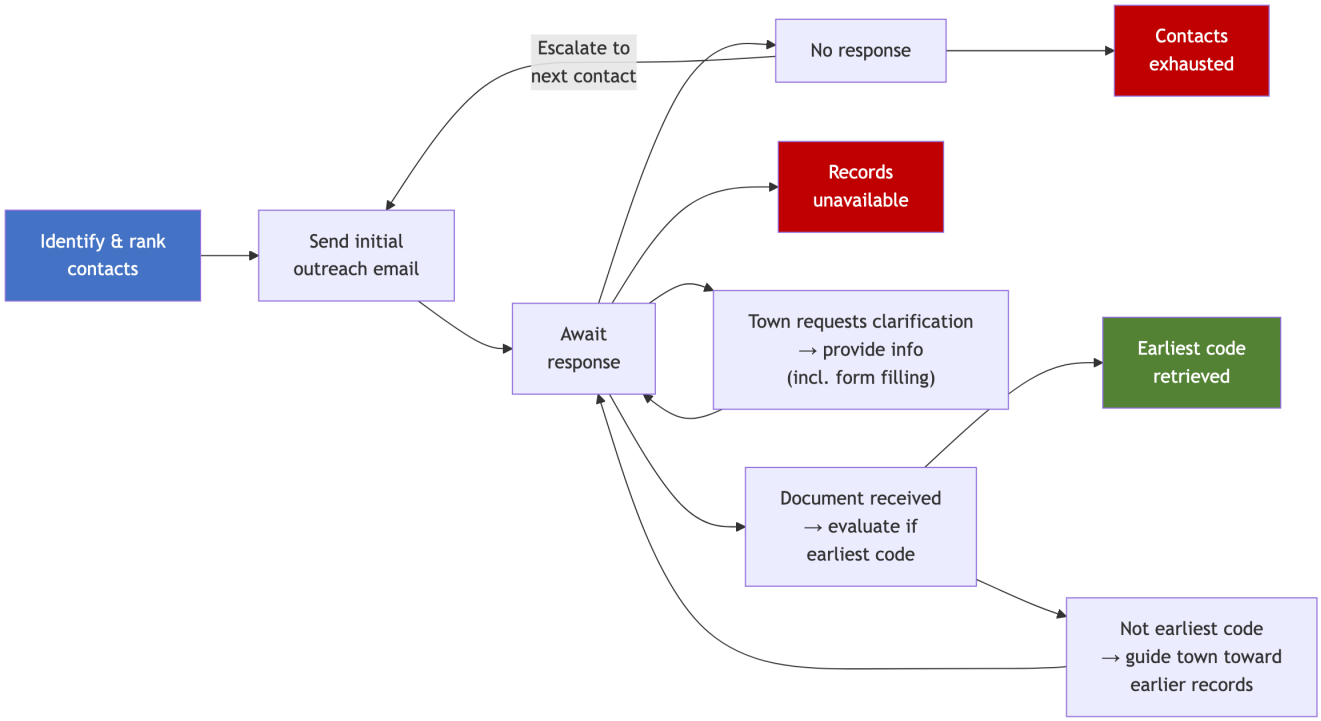
    - description: Ensuring adequate parking, street access, utility connections; subdivision design standards; coordinating private development with public infrastructure
    - keywords: parking, parking minimum, off-street parking, driveway, curb cut, street width, utilities, subdivision, sidewalk, access

9. `environmental_protection`

    - description: Protecting natural features and ecosystems; distinct from growth management's supply-restriction focus by targeting specific environmental resources
    - keywords: wetland, floodplain, steep slope, tree preservation, environmental review, watershed, stormwater, erosion, habitat protection"

# B Appendix Tables and Figures

Figure A1: Email outreach pipeline




Note: See Appendix section A.1 for more details.

Table A1: Multinomial Logit of Clause Type Category on Intent Category

| Outcome vs Procedural Intent Predictor | Bulk Restrictions | Land Use Restrictions | Miscellaneous |
|---|---|---|---|
| Aesthetic Design Control | 1.03 [0.84, 1.27] | 3.80*** [3.18, 4.54] | 18.00*** [14.29, 22.67] |
| Environmental Protection | 0.29*** [0.21, 0.41] | 5.36*** [4.16, 6.91] | 0.23*** [0.15, 0.35] |
| Exclusion | 2.18*** [1.93, 2.46] | 14.81*** [12.78, 17.16] | 0.19*** [0.12, 0.27] |
| Externality Control | 0.54*** [0.47, 0.62] | 27.02*** [23.55, 31.01] | 1.05 [0.85, 1.30] |
| Growth Management | 0.02*** [0.01, 0.02] | 0.71*** [0.57, 0.89] | 0.10*** [0.06, 0.15] |
| Infrastructure Coordination | 0.22*** [0.19, 0.25] | 1.65*** [1.44, 1.91] | 0.65*** [0.52, 0.80] |

*Notes:* This table contains a multinomial logit of a selected group of clause types on clause intents. Entries are relative risk ratios with 95% confidence intervals. The type outcome baseline is 'Procedural' and the intent predictor baseline is 'Public Health and Safety'. Standard errors are clustered by document.The sample includes all historical ordinances.

Figure A2: Initial outreach email template sent to municipal planning departments requesting historical zoning ordinances.

**Subject:** Historical zoning code request – *[Municipality], [State]*

Hi *[First Name]*,

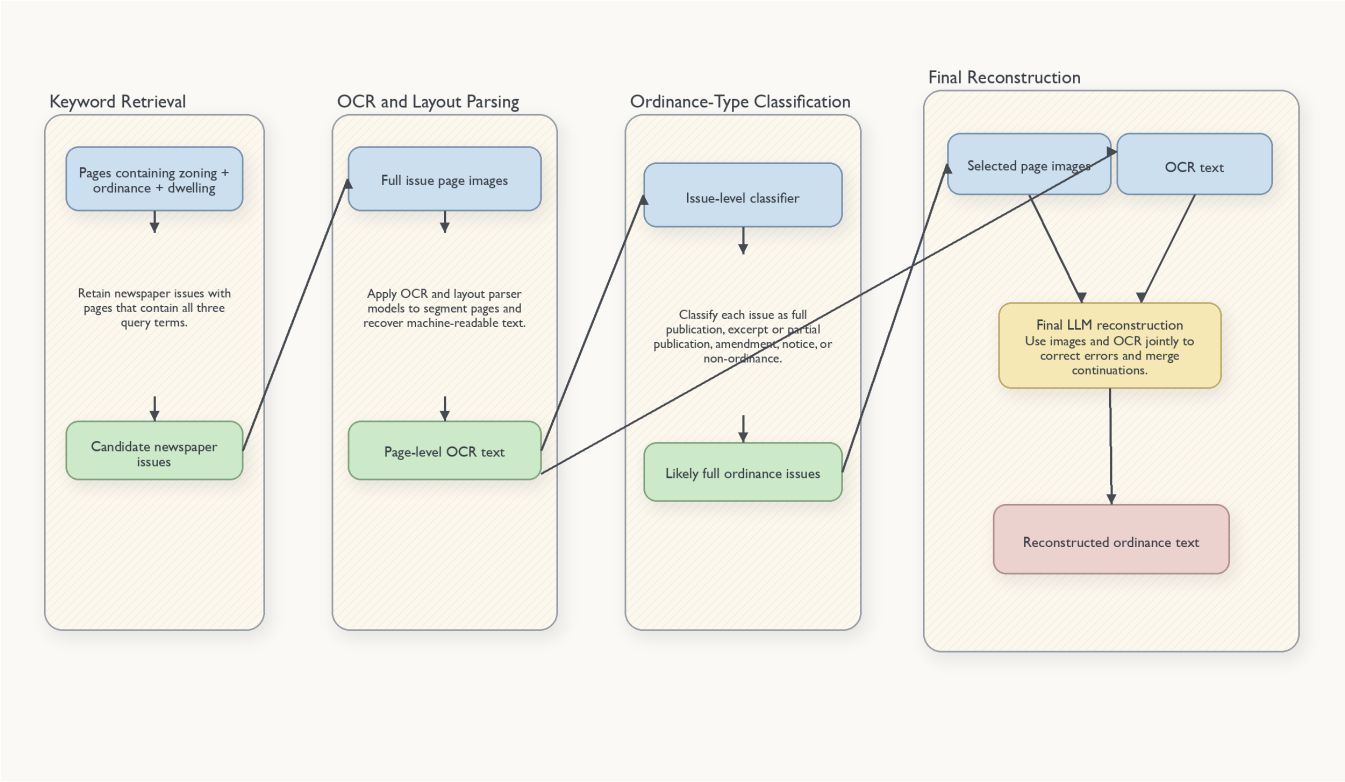
I'm Desmond O'Shaughnessy, a PhD student at New York University, and I'm compiling an archive of historical municipal zoning ordinances for an academic research project on land use and housing.

I'm specifically looking for the first zoning bylaw / ordinance your municipality enacted, along with the accompanying zoning map if you have it. If you can send me the earliest zoning code and map you have on record, along with the year of adoption, that would be a huge help.

If you're not the right contact for zoning records, could you send me the email address of the correct contact? If you'd prefer not to receive follow-ups, just let me know and I'll stop contacting you.

Thank you so much for your help,
Desmond O'Shaughnessy

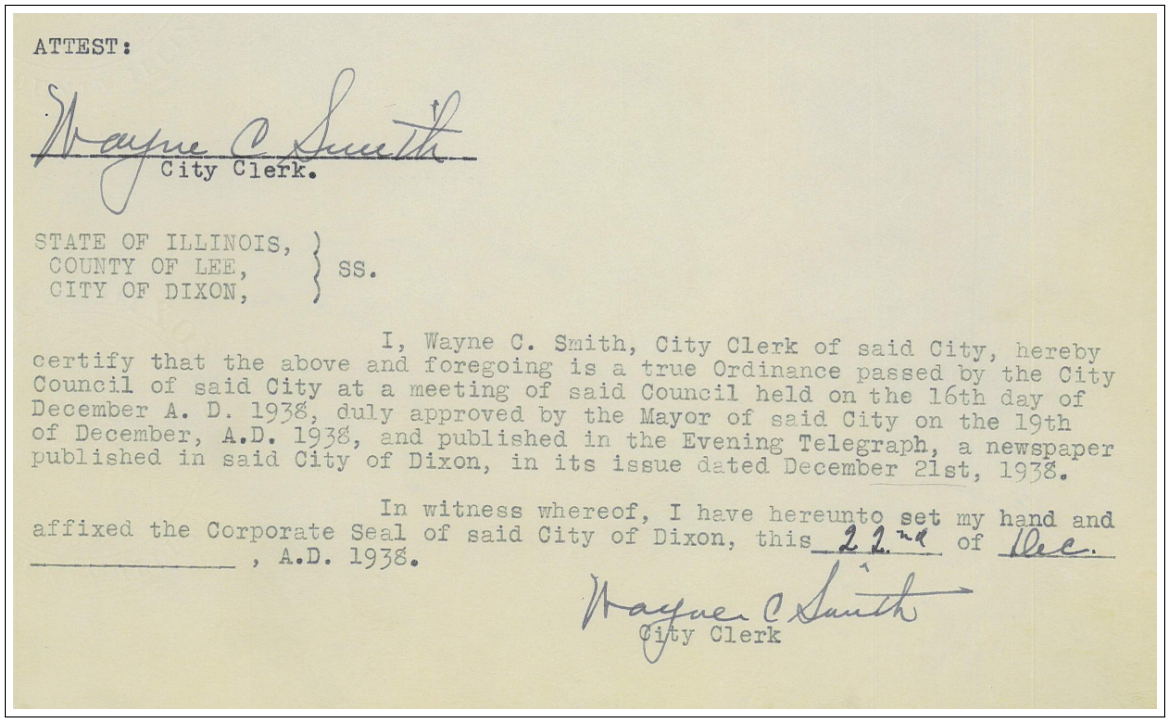
Figure A3: Historical Newspaper pipeline




Note: This diagram details the workflow that finds issues and pages containing zoning related keywords and uses them to construct full versions of municipal zoning ordinances.

Figure A4: Examples of newspaper publication references in zoning ordinances

(a) Dixon, Illinois (1938)



ATTEST:

Wayne C. Smith
City Clerk.

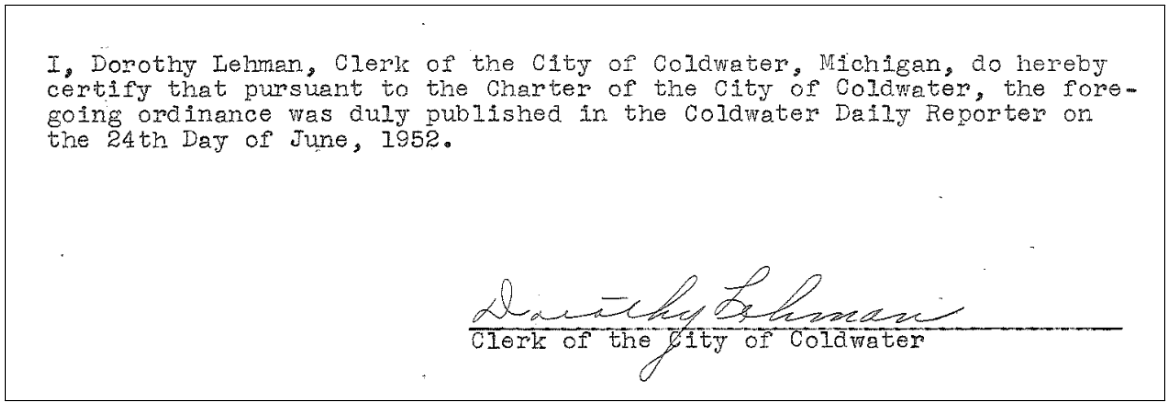
STATE OF ILLINOIS, )
COUNTY OF LEE, } SS.
CITY OF DIXON, )

I, Wayne C. Smith, City Clerk of said City, hereby certify that the above and foregoing is a true Ordinance passed by the City Council of said City at a meeting of said Council held on the 16th day of December A. D. 1938, duly approved by the Mayor of said City on the 19th of December, A.D. 1938, and published in the Evening Telegraph, a newspaper published in said City of Dixon, in its issue dated December 21st, 1938.

In witness whereof, I have hereunto set my hand and affixed the Corporate Seal of said City of Dixon, this 22nd of Dec. _____, A.D. 1938.

Wayne C Smith
City Clerk

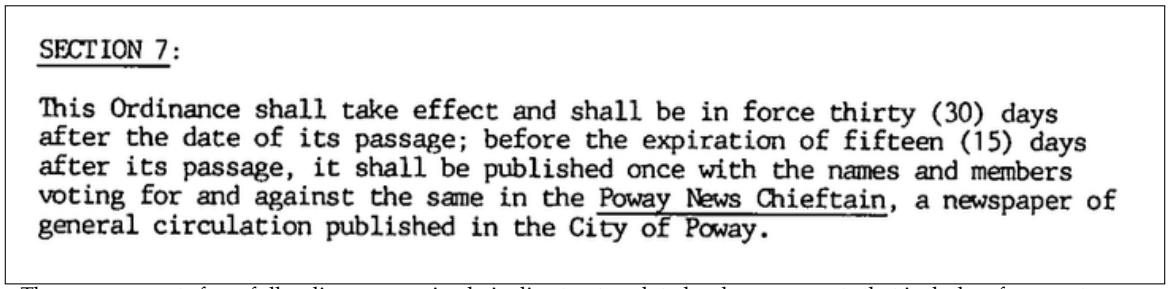
(b) Coldwater, Michigan (1952)



I, Dorothy Lehman, Clerk of the City of Coldwater, Michigan, do hereby certify that pursuant to the Charter of the City of Coldwater, the foregoing ordinance was duly published in the Coldwater Daily Reporter on the 24th Day of June, 1952.

Dorothy Lehman
Clerk of the City of Coldwater

(c) Poway, California (1983)



SECTION 7:

This Ordinance shall take effect and shall be in force thirty (30) days after the date of its passage; before the expiration of fifteen (15) days after its passage, it shall be published once with the names and members voting for and against the same in the Poway News Chieftain, a newspaper of general circulation published in the City of Poway.

Note: These are excerpts from full ordinances received via direct outreach to local governments that include references to newspaper publication.