The Political Economy of Moral Conflict: An Empirical Study of Learning and Law Enforcement under Prohibition*

Camilo García-Jimeno.†

JOB MARKET PAPER

Abstract

The U.S. Prohibition experience shows a remarkable policy reversal. In only 14 years, a drastic shift in public opinion necessitated two amendments of the U.S. Constitution. The adoption of many other policies and laws is similarly driven by initially optimistic beliefs about potential costs of their enforcement. Their implementation, in turn, affects the evolution of beliefs, giving rise to an endogenous feedback between preferences and policy choices. This paper uses data on U.S. cities during the Prohibition Era to investigate how changes in beliefs about the enforcement costs of Prohibition affected the mapping from moral views to policy outcomes, ultimately resulting in the repeal of Constitutional Prohibition. It first develops a dynamic equilibrium model in which communities make collective choices about law enforcement. Individuals differ in their baseline moral views about alcohol consumption and in their priors about the effects of Prohibition on crime. While both beliefs and moral views determine policy outcomes through the process of democratic decision-making, beliefs are in turn shaped by the outcomes of past policies. The model is estimated using a maximum likelihood approach on city-level data on public opinion, police enforcement, crime, and alcohol-related legislation. The estimated model can account for the variation in public opinion changes, and for the heterogeneous responses of enforcement and violence across cities. Shutting down the learning channel significantly limits the model's ability to match the moments of interest. The paper concludes with a series of counterfactual exercises that explore the equilibrium implications of changes in moral views, priors concerning the costs of enforcement, the degree of polarization in society, and the local political environment.

Very preliminary, all comments welcome

*I am extremely grateful to Daron Acemoglu, James Robinson, and James Snyder for their advice and guidance, and to Abhijit Banerjee and the participants at MIT's Political Economy Breakfast for their helpful suggestions. I also thank Maria Angelica Bautista, Angela Fonseca, Juanita Gonzalez, Catalina Herrera, Lucas Higuera, Thomas Morin, Michael Peters, and Maria Fernanda Rosales for their help at different stages of the project, and the staff at Harvard's Law Library Special Collections Room for their kind help. I also gratefully acknowledge the financial support of MIT's George and Obie Schultz Fund, and of the Banco de la Republica de Colombia. Errors are all my own.

†MIT, Department of Economics. 50 Memorial Drive, Of. E52-201, Cambridge, MA 02139. cgarcia@mit.edu.

"Man learns by the disappointment of expectations." Hayek (1960, p. 60)

1 Introduction

In *Individual Choice and Social Values,* Arrow (1963) argues that a proper understanding of collective choices requires taking into account the moral views of individuals because, as part of their preferences, they are analyticaly similar to externalities[1]. This insight proves particularly relevant in contemporary societies, where cultural heterogeneity is widespread and has been increasing over time, and polities are constituted by peoples with varying cultural backgrounds, and thus, different moral views. Indeed, differences in moral views have become a mayor source of disagreement about policy issues in many Western societies [2].

How differences in moral views affect policies is inexorably linked to individuals' beliefs about the implications of bans on certain activities, practices and expressions. While moral views and beliefs are mutually self-reinforcing, for example because those who find certain behaviors abhorrent also think that banning them can be effective and would have only minor unintended consequences, there is also a fundamental difference between moral views and beliefs. Moral views are slow-changing or even fixed, whereas beliefs about the implications of different types of bans and restrictions are frequently subject to a large extent of uncertainty, and can change rapidly as individuals observe their outcomes over time. Indeed, learning may be one reason why societies sometimes undergo radical social change and policy reform away from policies originally motivated by moral views, such as during the U.S. alcohol Prohibition experience of the early 20th Century. In this paper I argue that the reversal of Prohibition legislation in the United States can be understood as a result of belief changes about the implications of bans on the alcohol market. While Prohibition received support from a fraction of the population that held moral views against alcohol consumption, their beliefs that such bans could be implemented effectively and would reduce rather than increase crime contributed to their zeal. These beliefs changed rapidly, however, as communities experienced sharp increases in crime following the implementation of Prohibition. Many former supporters of the policy then found themselves in a situation similar to that of John D. Rockefeller, himself a radical prohibitionist, who recognized such a tension in the late 1920s:

> When Prohibition was introduced, I hoped that it would be widely supported by public opinion and the day would soon come when the evil effects of alcohol would be recognized. I have slowly and reluctantly come to believe that this has not been the result. Instead, drinking has generally increased; the speakeasy has replaced the saloon; a vast army of lawbreakers has appeared;

[1] *"From a formal point of view, one cannot distinguish between an individual's dislike for having his grounds ruined by factory smoke and his extreme distaste for the existence of heathenism in Central Africa... I merely want to emphasize here that we must look at the entire system of values, including values about values, in seeking for a truly general theory of social welfare." Arrow (1963, p. 18)* In Arrow's terms, an individual who performs a private activity which another individual considers immoral will, as a result, impose an externality onto him, out of the latter's regard of the former's action as immoral. Thus, for example, there is widespread agreement across individuals regarding the immorality of murder, but widespread disagreement regarding the morality of abortion.

[2] The salience of moral issues in the political agenda could be a result of convexity of preferences over them, as in Kamada and Kojima (2010), or because they are strategically exploited by an interest group, as in Baron (1994). In the context of Prohibition in the U.S., the latter seems a better description of the process leading to the adoption of Prohibition.

> many of our best citizens have openly ignored Prohibition; respect for the law has been greatly lessened; and crime has increased to a level never seen before. (John D. Rockefeller, quoted in Okrent (2003, p. 246-247))

In this paper I study the relationship between policy reform and social change, and argue that ex-ante uncertainty about the effects of radical changes in society's legal standards, coupled with the ability of individuals to learn about the effects of those policies, can be at the heart of the dynamics of social change, through a feedback between the effects of policies and changing attitudes in response to their effects, modulated by the endogenous extent of enforcement of those same policies. More specifically, I exploit the Prohibition experience of the 1910s-1930s to investigate the extent to which support for different types of bans is determined by the interplay between moral views and beliefs, and how this support changes as beliefs evolve as a result of learning from the outcomes of those policies.

In fact, as a methodological contribution, I argue that the mechanism proposed in this paper may have relevance outside the experience of Prohibition to understand the evolving attitudes towards moral issues, and more generally to think about the forces shaping social change. Attitudes towards Catholics in the 19th Century U.S., towards the role of women around the mid 20th Century, towards blacks in the South after the Civil War and after the Civil Rights Movement, or more recently towards Muslims in Western countries, for example, could be better understood by studying how the enforcement of policies targeted towards specific groups has effects that change collective preferences over those policies, endogenously feeding back into changes in policy choices, and in individual attitudes in the long run.

With this purpose, I develop and estimate a dynamic structural model of Prohibition enforcement and crime, where heterogeneity in moral views and beliefs interplay, and have observable and unobservable components. Learning is rational, and communities decide the enforcement margin of Prohibition through a collective decision. Law enforcement shifts the distribution of crime, and individuals update their beliefs about the effects of Prohibition by observing homicide rate realizations. Because law enforcement is endogenous to preferences and beliefs, the speed of learning by rational agents is affected not only by their priors, but also, indirectly, by the distribution of moral views giving rise to such collective choices of law enforcement.

I estimate this model by Conditional Maximum Likelihood, using a dataset of U.S. cities during the period 1911-1936, when the country experienced a Prohibitionist wave which reached Constitutional status, and focus on the homicide rate, the drunkenness arrest rate, and police expenditure as the main observable outcomes. I start by showing that crime and law enforcement during Prohibition presented a rise and fall pattern, and that the alcohol market contracted and rebounded quickly thereafter (see figures 1-2). Then I document how these patterns differed between cities with varying moral preferences, by using observable variation in the distribution of religious ascriptions and other demographics: drier (i.e., more favorable to Prohibition) cities experienced initially higher levels of law enforcement, while wet (i.e., less favorable to Prohibition) communities observed higher increases in criminality and larger changes in public support for the policy.

The estimated structural model explains a large fraction of the variation, both across cities, and over time, in the choices of policing expenditure in cities, the observed evolution of criminality measured through the homicide rate, and the alcohol-market dynamics. With the model I also estimate the extent to which Prohibition as a legal standard, and Prohibition enforcement, were responsible for the increase in criminality observed during the period. Prohibition was associated with an average homicide rate increase of 15 to 20%, while it was unable to shrink the alcohol market. At its lowest point, around three years into Prohibition, the effective alcohol supply fell by around 35%, but rebounded quickly thereafter. Moreover, I estimate that the Prohibition-related homicide rate was increasing in the level of law enforcement. Relatedly, cities in smuggling areas had a lower potential for crime to develop under Prohibition, and I argue this was due to the reduced constraints faced by the black alcohol market in those areas.

The structural model also allows for the estimation of several moments of the joint distribution of moral views about Prohibition, and prior beliefs about its effects. I find that beliefs were extremely optimistic across the distribution of moral views, so that the variation in moral views across cities was larger than the variation in initial beliefs. Although people had strong opinions about alcohol Prohibition at its outset, there was not much disagreement about its effects. Nevertheless, the estimated correlation between moral views and beliefs is large, implying that drier individuals held even more optimistic prior beliefs about the effects of the policy.

I conclude with a series of counterfactual exercises based on the structural model, which illuminate the key interactions taking place during Prohibition. I find that local policy was highly responsive to community preference changes. As a result, a more polarized society would have learned faster, but also would have observed higher crime increases during Prohibition. Communities would have responded to Prohibition by offsetting it with reduced law enforcement choices if prior beleifs had been less optimistic; this would have reduced the crime spike of the 1920s, but would have limited the speed of change in public opinion. Finally, in an exercise where local decision-making power is shifted away from the median voter, the increased misalignment between the community's distribution of preferences and the equilibrium law enforcement choice alters the speed of learning by changing the informativeness of the crime signals.

1.1 Related Literature

This paper is related to several research areas. The first studies the determinants of civil liberties. To my knowledge, Lagunoff (2001) is the only work which directly addresses the question of why democracies are able to sustain civil liberties for minorities. According to his argument, when a majority is likely to become a minority in the future, it will have incentives in the present to weaken the enforcement technologies available, which could otherwise be used against them in the future. Political Scientists, on the other hand, have stressed that the salience of moral issues is relevant to understand the extent of civil liberties, because it determines the degree to which the legal standards adopted will respond to interest group politics (Haider-Markel and Meier (1996)). Indeed, through the political system, different practices are prohibited or restricted based on moral motivations alone. In autocractic societies, rulers and elites directly impose their moral views upon the community; in democracies, majorities can impose restrictive legal standards upon minorities through the ballot box. Legal restrictions on individual liberties are of economic and political importance for several reasons. First, they directly have welfare implications over both individuals who favor and disfavor the prohibition. More interestingly, they often have potentially uncertain side effects. The imposition of a restrictive legal standard creates dissatisfaction in a subset of the population, leading to non-conformist behavior, political mobilization, unrest, or violence.

As a result,*de jure* prohibitions require concomitant *de facto* enforcement. Because they are prone to widespread loopholes, enforcing restrictions on the behavior of individuals requires costly monitoring and willingness to enforce, both by the community and by its enforcement agents. In fact, within the literature on "crime and punishment" pioneered by Becker (1968), concerned with the understanding of the determinants of crime enforcement and the effects of law enforcement on the equilibrium levels of illegal activity, this paper highlights that what society defines as crime is endogenous, and that, as a result, punishment is a social chioce. These considerations have been overlooked in the literature, and suggest that agreement about punishment within society, and social learning about its costs and benefits, might be important to understand the success of alternative policies.

In this paper, the main channel driving public opinion and law enforcement outcomes is the interaction between beliefs and moral views, making it close to the research on policy and rational learning. Landier et al. (2008) and Alesina and Fuchs-Schundeln (2007) study how ideological differences have affected beliefs about capitalism. Sargent et al. (2006) develop a statistical model about monetary policy in the U.S., where policymakers endogenously learn about the Phillips curve. Buera et al. (2010) is a recent example of structural estimation of a learning model, where policymakers update their beliefs about the merits of market oriented versus interventionist policies by observing their neighboring countries' outcomes. A theoretical paper in the same spirit is Mukand and Rodrik (2005), who argue that experimentation and imitation might explain why, over the last decades, countries have converged in the adoption of policies, but not in economic performance. Strulovici (2010) is also an important recent contribution, which studies the incentives for policy experimentation in a dynamic voting framework, in which incentives for experimentation are limited by the trade-off between learning about the effects of policies and the pivotality of voters.

Finally, this paper also contributes to the literature on crime (See Dills et al. (2008) for a recent survey where the authors conclude that the most robust correlate of crime is the prohibition of drugs). The literature has stressed factors like the age composition of the population (and abortion), the deterrence effects of incarceration, access to firearms, investment in policing, inequality, or the economic cycle (Levitt (2004); Donohue and Levitt (2001); Dills et al. (2008)). Miron (1999a) and Goldstein (1985) stress the main channel I explore in this paper, where non-conformism and law enforcement over activities involving traded commodities create the potential for violence and corruption to arise as salient side effects. Competition for black market rents, unavailability of institutionalized channels of dispute resolution, and the use of coercion by law enforcers, all create incentives for crime and corruption to develop (Miron (1999b)). Thus, tightening law enforcement can magnify the effects of the prohibition on crime. It drives out the marginal producers (which are less likely to engage in criminal behavior), weakens social norms sustaining peaceful dispute resolution among criminals, and crowds out resources for overall crime enforcement (Becker (1968); Miron (1999b)). The literature has mostly focused on reduced-form or instrumental variables strategies, whereas I explicitly model the endogenous relationship between law enforcement choices and crime that arises in the context of Prohibition, highlighting the role of rational learning and beliefs.

The rest of this paper is organized as follows. Section 2 presents a historical overview of the Prohibition experience in the United States during the early decades of the Twentieth century, and discusses its institutional and political background. Section 3 then presents and discusses the data collected and used in the paper. Based on the historical discussion, section 4 subsequently presents reduced-form results, which guide the development of the model presented in section 5. Section 6 proceeds with the estimation results from the structural model, and presents some counterfactual exercises. Finally, section 7 concludes.

2 Prohibition: A Historical Overview

2.1 Prohibition Politics

Nation-wide alcohol Prohibition in the United States was written into the Constitution as the 18th Amendment in January 1919, and repealed from it just fourteen years later, as the 21st Amendment, in December 1933. Given the constitutional supermajority requirements to amend the U.S. Constitution, such a policy reversal is striking[3]. The increase in criminality during the Prohibition period, best illustrated by figure 1, was as striking, and a first-order reason why public opinion had such a radical swing in such a short period of time. Alcohol Prohibition, though, was not a sudden appearance; it was the endpoint of a prohibitionist wave with origins dating as far back as the 1870s, when a group of Ohio women organized the so-called Temperance Crusade, which would later give rise to the Women's Christian Temperance Union (WCTU).

Prohibition was introduced staggeredly across counties and states through a gradualist political strategy of religiously motivated temperance groups, closely related to the Baptist, Methodist and Evangelical churches, and composed mostly of native-born whites and women (Sinclair (n.d.); Okrent (2010)). The two most prominent were the WCTU, and the Anti-Saloon League (ASL). Both developed a nationwide organizational structure, but the ASL took the lead in the beginning of the Twentieth century. Initially these groups were not a majority of the population. Their political success was due to their pivotal character in the competitive context of bipartisan politics, based on strong campaigning and lobbying in state legislatures, towns, and cities, and on the intensive use of referenda initiatives. Republicans and Democrats were frequently so evenly divided that a switch of the temperance vote could easily decide local elections. Prohibitionist groups were able to become pivotal even in the within party races of the Democratic-dominated South. Their persistence in lobbying also was important because Prohibition was not an issue that politicians paid much attention to at the time[4].

Key to the political success of the drys was their strategic avoidance of aligning with either party. While the ASL was relatively antagonistic to Northern Democrats whose constituencies were mostly in large urban areas, it was much closer to Southern Democrats for whom Prohibition was another

[3]Constitutional amendments require approval by two thirds of the vote in both the House and the Senate, and a plurality of the vote in either both chambers of at least three fourths of the State Legislatures, or in at least three fourths of State Constitutional Conventions.

[4]Talking about the 18th Amendment, Sinclair (n.d., p. 182) argues that "... boredom played some part in the passage of the amendment. The members of Congress were sick of being badgered by the Anti-Saloon League and their dry constituents."

channel for social and political control of blacks (Asbury (1950, p. 93), Sinclair (n.d., p. 182)). There was disagreement on the issue within Democrats in the South too, as a faction of the party believed that allowing the Federal government to make decisions regarding Prohibition could be the first step to further undermine Southern autonomy (Szymansky (2003)). An indicator of the lack of partisan alignment on Prohibition is the House roll call on the 18th Amendment; 64 Democrats and 62 Republicans voted against, while 140 Democrats and 138 Republicans did so for Prohibition. A second important element to explain the success of the dry campaign was its gradual approach. Local option measures were followed by state-wide legislation, so that right before the 18th Amendment was adopted, almost 80% of U.S. counties were already under some form of Prohibition, starting from no more than 15% in 1900. Figure 3 shows the dates of state-adoption of Prohibition legislation.

The rise of prohibitionist attitudes in the U.S. was part of the so called "Progressive Era", a much broader set of social and political changes taking place in the late Nineteenth and early Twentieth Centuries, associated with the rapid expansion of State capacity throughout the country. The establishment of the income tax under the 16th Amendment and the enfranchisement of women under the 19th Amendment also were part of the expansion of the role of the State in society. In this perspective, Prohibition expanded the role of the State into the private activities of individuals. This required an unprecedented involvement of the churches in politics, fueled by a context of rapid social change and urbanization, which was increasing the heterogeneity of the American society. On the dry side, priests moved from claiming the sinfulness of drinking[5], to advocating explicitly prohibition legislation (Isaac (1965, p. 263)). On the wet side, it was about "whether or not the American people were going to hand over to government the paternalistic power to regulate lives and habits" (Kyvig (1979, p. 51)).

2.2 Law Enforcement

Before Constitutonal Prohibition, enforcement of the alcohol laws in states under Prohibition was usually mild. In dry communities it was redundant, while in wet communities it was relatively ignored. A large share of alcohol consumption took place in saloons and other public spaces, which made public intoxication a widespread phenomenon (See, for example, Blocker (2006); Stayton (1923)). Prohibitionist associations were concerned about the social consequences of saloons, and arrests for drunkenness were seen as a key indicator of successful enforcement of dry laws. But loopholes were abundant and often overlooked (Franklin (1971)). The biggest loophole was probably interstate shipping of alcohol into states under Prohibition. As a response, the ASL lobbied intensively until it achieved the passage of the Webb-Kenyon Act in 1913, banning interstate shipping of alcohol into dry states. Although later it was practically unenforced, at the time of its passage this law was very controversial. President Taft vetoed it, and Congress subsequently overrode his veto. Wets, backed by the Brewers Association, argued the law violated the First Article of the U.S. Constitution, but the Supreme Court later upheld it.

[5] Okrent (2010, p. 33) quotes a WCTU strategist who, being asked why alcohol was inconvenient, gave the following account: "...selling in prohibited hours, gambling, selling to intoxicated men, rear rooms, unclean places, invading residential districts, the country saloon, the social evil, selling to minors, keeping open at night, brewers financing ignorant foreigners who are not citizens, the American bar, brewery-controlled saloons, cabarets, Sunday selling, treating, free lunch, sales to speakeasies, bucket trade, signs, screens, character of the men, too many saloons".

At the same time, although the passage of the 18th Amendment and its enforcement law (the Volstead Act)[6] appeared as highly restrictive by banning any liquor with more than 0.5% alcoholic content, Congress did not make large appropriations for its federal enforcement. In fact, the Amendment established concomitant enforcement by the local, state and federal levels, so Congress, expecting cooperation from local and state policing agencies and general compliance with the law, created a modest federal enforcement organization (Kyvig (1979, p. 23)). The weakness of federal enforcement is best exemplified by the constant changes in Prohibition administration during the 1920s [7].

Table 1 presents the main Federal Prohibition law enforcement outcomes during the 1920s. Trends are very similar across the four main U.S. regions, and suggest that enforcement intensity peaked around 1928. Early during national Prohibition, given the initial absence of domestic producers, most of the supply of illegal liquor came from international smuggling (Okrent (2010)). Over time, local production based on illegal distilleries and stills caught-up with demand. Nevertheless, the number of distillers and fermenters seized fell sharply in the later Prohibition years, which suggests a sharp fall in the enforcement activities against producers. The number of killed or injured agents during enforcement activities shows that the 1927-1930 period was particularly violent, and that subsequently few risky law enforcement activities took place. The 1927-1930 period coincides with the years in which Prohibition administration was under the Bureau of Prohibition, in what historians have acknowledged as the last attempt form the federal government to control the liquor trade, in response to the fall in state and local law enforcement throughout the country. Indeed, by 1928 several states had already repealed their own enforcement legislation. To have an idea of the limited extent of law enforcement at the federal level, notice that in 1929-1930, total liquor seizures in the U.S., including spirits, malts, wines, cider, mash, and pomace, were approximately 74 million gallons. Compared to the $3,375$ million gallons of booze which, according to Okrent (2010, p. 202), were produced and distributed annually by Max Hoff, an illegal producer in Pennsylvania, the federal enforcement looks almost irrelevant.

In fact, most of the law enforcement, in practice, relied on local efforts. This was not only because of the inherent difficulties in enforcing alcohol restrictions throughout the country, which limited the federal law enforcement stategies to infrequent raids and a focus on some particularly troublesome areas, but also because of the inefficiency of the federal agency. Complaining about this issue in 1926, Colvin (1926, p. 497) argued that, "Although the United States had adopted a national standard throughout the nation, the administration of the law so perverted this objective as to make enforcement substantially a matter of local opinion because it was administered to so large a degree by men owing their appointment to local political influences and subject to local political

[6] President Wilson also vetoed the Volstead Act, and his veto was also overridden by Congress.

[7] Originally, the Volstead Act created the Prohibition Unit as a department of the Bureau of Internal Revenue, with Prohibition Directors in each state. The Coolidge administration avoided dealing with the Prohibition problem throughout, and in 1925, there was a sharp reduction in the size of the Prohibition Unit (Colvin (1926, p. 495)). The critical situation regarding corruption and venality within it resulted in a reform of Federal Prohibition administration under the Prohibition Reorganization Act of 1927. This act created the Bureau of Prohibition, ascribed to the Treasury Department, putting its employees under the Civil Service and creating 27 Prohibition Districts (Schmeckebier (1929), Schmekebier (1923)). Finally, in 1930 the Prohibition Bureau was transferred to the Justice Department, but at this point, "...as useful as these congressional steps may have been... the enforcement effort had acquired a dismal reputation and doubts as to whether Prohibition could possibly be effective had become deeply engrained" (Kyvig (1979, p. 32)).

pressures... it was the worst form of local option -the option of the local politicians to determine the extent to which the law should be enforced-, politicians, many of whom were personally wet, others of whom wanted to placate a wet element in their constituencies, and all of whom belonged to political parties which sought wet votes as well as dry ones". While a dry such as Colvin saw the problem in the ineptitude and corruption of enforcers, a wet such as Tydings would argue that "If moral force... does not make them stop, physical force must be used. Where is the physical force to come from? Plainly, in a nation of 120 million people, scattered over an area of 3 millon square miles, the force must be predominantly supplied by the local enforcement authorities... but the police, the courts and the juries are the servants and reflectors of local sentiment"(Tydings (1930, p. 125)).

Thus, the degree of law enforcement of Prohibition was responsive to the local demand for both Prohibition and alcohol, and elected authorities were agents of both groups. This seems to have been true not only during Constitutional Prohibition, but also during state-level Prohibition. Franklin (1971), for example, quotes a local judge in dry Oklahoma claiming that a candidate for sheriff would not possibly be elected, if it were known that he intended to enforce Prohibition. In the same way, judges and juries tended to be lenient in their decisions regarding Prohibition violation cases (Szymansky (2003, p. 184), Kyvig (1979, p. 25), Tydings (1930, p. 127)). Judicial leniency was even institutionalized through the so-called "bargain days",[8] which arose in response to the courts' congestion created by the overwhelming number of violations of the Volstead Act. In fact, initiated criminal prosecutions in federal courts for violations of Prohibition increased from slightly more that 100 per million inhabitants in 1920, to almost 500 in 1925, which made up 80% of all criminal prosecutions[9].

If law enforcement varied as a function of local preferences, the effects of Prohibition also varied between communities. This is acknowledged by a Commissioner traveling around the State of New York in 1930 who argued that the problems varied between and within states, particularly between the rural and urban areas[10]. According to Kyvig (1979), Scandinavians in Minnesota continued to drink, while Idaho, Oregon, and Washington had come to accept Prohibition. Los Angeles and even San Francisco had large dry constituencies, and relatively dry areas ran from California to Texas. Louisiana, on the other hand, was extremely wet and law enforcement relied almost exclusively on federal authorities. In the rest of the South, Prohibition was enforced particularly on blacks. Finally, in the large wet cities of the Northeast such as Detroit, Cleveland, Pittsburgh, Boston and New York, Prohibition was largely unobserved, and weakly enforced, particularly after the second half of the 1920s.

The weakening of law enforcement took place not only by a reduction in policing and prosecution, but also through the repeal of state enforcement legislation. The most prominent case was that of New York, which very early on, in 1923, repealed the state enforcement law. Alfred Smith, the

[8] Violators would plead guilty and be charged a small fine.

[9] I collected the data on judicial prosecutions at the judicial district level for the period 1915-1933 directly from the Attorney General Annual Reports.

[10] "New York City presents a problem quite distinct from the up-state section, and the border region presents an entirely different situation... the problem varies as the population is homogeneous or heterogeneous... throughout the rural and smaller cities... there is a greater respect for the law and established order" (Wickersham-Commission (1928-1931b, Box 13-2, Prohibition Survey of New York, p.2)).

Democratic Presidential candidate in the 1928 election, was then the Governor of New York. The repeal was by no means a consensual decision, and in fact, many dry organizations lobbied Smith to veto the it. It was a difficult decision because, although openly wet, alienating the dry vote could prove costly for his future political career. In his own words, "Some seem to think that my approval [of the repeal] will mean the preservation of American Institutions. Many others impeled by equally patriotic motives seem to feel that my approval will be destructive of American government. Obviously, both cannot be right..." (Smith (1923, p. 601)).

2.3 Repeal

The early repeal of state enforcement legislation in New York was driven more by the morally anti-Prohibitionist character of its large share of urban population than by a rise in criminality, which by that time, had not yet peaked. The shift in public opinion in other regions of the country took place at a slower pace, and more in response to the observable increase in criminality. Initially dry individuals, who were morally compeled by Prohibitionist reasoning, could not avoid acknowledging the adverse consequences that the policy was having.

The rise in crime and undermining of the rule of law was not homogeneous across the country, and as a result, neither was the fall in support for the policy. The Democratic party, which had been out of power throughout the 1920s, managed to capture most of the rise in anti-Prohibitionist sentiment. In the 1928 Presidential election this hurt Al Smith, but in the 1932 campaign it played in favor of Franklin D. Roosevelt. The distribution of public opinion did shift massively against Constitutional Prohibition, and opposition became better organized. The Association Against the Prohibition Amendment, for example, began its advertising campaigns in 1928, focusing on providing information about the ill-effects of Prohibition. In 1929, the Women's Organization for National Prohibition Reform was founded with the same intentions. Nevertheless, even after the repeal of the 18th Amendment, six states remained dry[11]. Among the rest of the states, some instituted systems of "state operation", in which the state directly controlled the distribution of alcohol, while others just imposed some regulation over a free market (Harrison and Laine (1936, p. 43)).

3 Data and Summary Statistics

3.0.1 Data on Criminality

Criminality was the main source of concern and learning about Prohibition for the public. The homicide rate is the variable for which most comprehensive information is available, and one for which measurement error is likely to be very limited. Thus, I collected information from the *Mortality Statistics* published yearly by the Bureau of the Census, reporting the number of non-traffic-related homicides for a sample of U.S. cities. I complemented this information with the homicide data reported in the Wickersham Commission documents, finally putting together yearly data for the

[11] These were Alabama, Kansas, Mississippi, North Carolina, North Dakota and Oklahoma. Nonetheless, all of these, except Alabama and Kansas, allowed for the sale of beer (Kyvig (1979, p. 188))

period 1911-1936 and a sample of up to 93 cities. Data on drunkenness arrests, on the other hand, is very detailed and covers a total of 573 cities for the period 1910-1929[12]. Finally, the Federal Bureau of Investigation began compiling and publishing its *Uniform Crime Report* (UCR) in 1930, which contains yearly city-level data on murders and other offences reported to the authorities. Offences include robbery, assult, burglary, larceny and auto theft.

3.0.2 Law Enforcement Data

Law enforcement is a difficult concept to measure because it depends on the discretion of the enforcer, and thus, is necessarily unobservable. Moreover, measuring law enforcement through its outcomes is problematic; an increase in liquor stills seized, for example, could be explained by an increase in Prohibition enforcement on a constant level of illegal alcohol production, or by a reduced level of law enforcement which allows for illegal production to increase. Because a great deal of Prohibition enforcement, and all of local crime enforcement, was decided and implemented at the city level, I focused on collecting data on city public finances, and specifically, on police expenditure. I use the *Financial Statistics of Cities* published yearly by the Bureau of the Census, which report disaggregated data on city public finances for cities with populations above $30,000$ (around 250 cities), and obtain data on total city public expenditure and investment, police expenditure and investment, and all protection expenditure and investment (all protection includes police, fire and other expenditure), for the period 1911-1936. I computed 1913-constant prices expenditure data by using the U.S.-wide CPI as of June of each year as the deflator[13].

3.0.3 Demographic and Religious Data

City and county-level data on demographic characteristics are taken from the decennial population censuses. I focus on the age distribution, the ethnicity distribution[14], and total population, from the 1910-1940 Censuses. Given the strong relationship between ethnicity and religiosity with attitudes towards the liquor problem, I use religious ascription data from the decennial Censuses of Religions (1906, 1916, 1926, and 1936), to capture heterogeneity in moral views about Prohibition. I aggregated religious ascriptions in the following nine groups, directly from their names: Baptist, Eastern Orthodox, Evangelical, Jewish, Mormon, Lutheran, Methodist/Episcopal, Catholic, Presbyterian, and other. The consensus amongst historians is that Baptist, Evangelical, Mormon, Methodist, Episcopal and Presbyterian communities held the strongest views in favor of Prohibition, while Catholic, Orthodox, Jewish and Lutherans had much more favorable positions regarding alcohol consumption (See Foster (2002); Lewis (NA); Szymansky (2003)). I refer to the former as “dry”, and to the latter as “wet” religions. I then computed the share in each religion directly as the number of adherents divided by the total number of adherents to any religion in the city (or county).

[12] The data on drunkenness arrests contained in the Wickersham Commission papers appears to have been originally compiled by the World League. Dills et al. (2005) use this source, covering a shorter time period, together with an alternative source compiled independently by the Moderation League. Both series appear to be highly correlated, so I restrict attention to the World League data, which covers the whole 1911-1929 period.

[13] Data for the years 1914 and 1920 is unavailable. For the balanced panel estimations below, I use the interpolated values (1913-1915 average for 1914, and 1919-1921 average for 1920) for these two years.

[14] I focus on the distribution of the population between native white, foreign white, and black individuals.

3.0.4 Public Opinion Data

To measure public opinion about Prohibition, I collected electoral returns data on referenda on alcohol-related issues for the different states, taking place during the 1900s-1930s. These referenda were usually ballot measures proposed to the citizens to approve or repeal liquor laws, or ammend the state constitutions. In states where local option was in place, county or city-level referenda had the purpose of allowing or forbidding the sale of alcohol. When submitting the 21st Amendment to the states, the U.S. Congress determined that Constitutional Conventions should be elected in the different states to decide over the issue, and candidates should run in either a dry or a wet slate (Brown (1935)). All of the referenda returns allow me to directly compute the fraction of (anti-Prohibitionist) wet vote, which I use as a proxy of wet support[15]. Almost all of the electoral returns data is available at the county level, except for referenda in the states of Connecticut and Massachusetts, for which city-level data was reported. Overall, I have referenda election returns for $2,083$ counties.

3.0.5 Legislation Data

Alcohol-related legislation across states comes from three main sources: the Anti-Saloon League's 1916 Yearbook and the information in Szymansky (2003), and in Cherrington (1920). The latter source was in particular very useful since the author makes a state-by-state compilation of all of the dry legislation up to 1920, detailing the time of its passage and/or repeal, and providing a brief description of it. Based on these sources, I coded a state-level variable for the number of dry laws in place in each year, an indicator variable for being under Prohibition (either state-level or federal-level), and an indicator variable for having a Prohibition enforcement law in place.

3.1 Summary Statistics

Table 2 reports population-weighted summary statistics for the main variables used in the paper, summarizing the available information for up to 340 cities (counties for the referenda election returns data), and disaggregating the sample in the four main U.S. geographic regions. The table presents the baseline distribution of religious ascriptions and demographics, together with data on legislation. It also includes summary statistics for the different outcomes of interest, comparing average values in the 1910s and 1920s.

For the religious distribution, I present summary statistics from the 1916 Census of Religions. As expected, Southern cities were heavily Baptist and Methodist relative to the rest of the country (29% and 24% respectively). The South was also less Lutheran and Catholic. Indeed, Catholicism was concentrated in the Northeast and Midwest, where more than half the adherents in the sample belong to this religion. Evangelicals were mostly concentrated in the Midwest, while Mormon communities were mostly found in the West. In fact, with almost a 50-50 split between dry and wet religions,

[15] The main caveat here is that differences in turnout rates might differ systematically between Prohibitionist and anti-Prohibitionist voters, not reflecting the true distribution of political preferences in the community. For an empirical model of turnout on alcohol-related referenda, see Coate and Conlin (2004).

the Western cities present the more uniform distribution of religious membership . In contrast, religious membership in Southern cities was heavily skewed towards dryness, while in the Midwest and Northeast wet religions were majoritarian.

Looking at the basic ethnic composition across regions from the 1910 Population Census, 26% of the population in the Southern cities in the sample was black, in sharp contrast with all other regions where the black population was between 1.3 and 3.1 percent. The foreign white population was especially prevalent in the Northeast, where 32% were whites born outside the United States, as compared to only 7% in the South. In the Midwest, on the other hand, almost three quarters of the population was native white.

A look at the outcome variables reveals that real per capita expenditure in police was significantly larger in the 1920s than in the 1910s, with an average increase of around 0.3 dollars. Northeastern cities had the highest levels of expenditure in both decades, but Southern cities experienced the largest average increase. Although per capita expenditure in police rose, the data on police expenditure as a share of total city expenditure reveals a fall everywhere, due to the fast increase in public spending in other categories during these Progressive Era decades. Cities in the West had the lowest police shares (around 8%). While per capita policing was lowest in the South, Southern cities had the highest share of their budget allocated to police ($11-12\%$). The average behavior of the data on drunkenness arrests reveals considerable differences between regions. In Southern cities, average arrests were very similar in the 1910s and 1920s. In contrast, cities in the West do show a large fall in arrests for drunkenness between both decades, falling from 22.5 to 13.9 per $1,000$ inhabitants. Although arrests in the Midwest and Northeast also are somewhat lower in the 1920s, the fall is not as large.

The homicide rate, on the other hand, shows significantly higher levels in the 1920s in all regions, and large level differences across them. While homicide rates were on average 5.3 per $100,000$ in Northeastern cities during the 1910s, they were almost five times higher in the South during the same decade. The variance of the homicide rate was also much larger in the South. It is also worth noticing that the smallest average increases in the homicide rate took place in the West, where it only increased from 9.8 to 11.6.

Support for Prohibition, as measured by the electoral returns on alcohol referenda, was higher in the South and the West, where the wet vote shares were 0.46 and 0.45 on average, while it was slightly above 50% in the Midwest and the Northeast. A comparison of these numbers between decades reveals the striking shift in public opinion; wet support was around 20 percentage points higher in the West and Midwest, 30 percentage points higher in the Northeast, and 10 percentage points higher in the South after Prohibition. Interestingly, the South showed the smallest increase in wet support, while, despite its higher initial anti-Prohibitionism, Northeastern cities experienced the largest average shift against Prohibition.

4 Some Reduced Form Results

I begin the empirical analysis by focusing on three first-order sources of variation in the effects of Prohibition. Differences in the timing of its adoption across states, in preferences over the legal standard (i.e. moral views about alcohol consumption), and in state-level legislation and its enforcement both at the local and federal levels. I focus on three outcome variables: as a direct measure of criminality, the homicide rate shows a large increase, happening with some delay after the introduction of Prohibition, and reaching its highest levels around the mid 1920s. Crime increases were larger in cities with bigger potential alcohol markets and populations less inclined to the policy. However, they were similar between cities facing different initial crime levels. By looking at the drunkenness arrest rate, I document a drastic contraction of the alcohol market right after the introduction of Prohibition, but a steady and relatively fast recovery. Neighboring markets reduced the extent of contraction in alcohol consumption, and the time path was remarkably similar across different cities. Finally, there is evidence of a steady increase in law enforcement following the introduction of Prohibition, with a subsequent fall starting in the late 1920s. The early increases in law enforcement were faster in cities with constituencies more favorable to Prohibition, but for late Prohibition years, these cities show lower spending in policing. Subsequently, I look at changes in public opinion regarding Prohibition by exploiting electoral data on liquor referenda, and document a non-monotonic relationship between changes in public opinion and overall moral views of cities' constituencies: communities with intermediate levels of initial support towards the policy saw the largest shifts in public opinion against Prohibition.

4.1 Crime, Law Enforcement, and the Timing of Prohibition

A natural first approach is to compare outcomes before, during, and after repeal of Constitutional Prohibition. Figure 1 shows that the advent of Prohibition saw a sharp increase in crime, here measured by the homicide rate (although a mild, positive pre-trend can be observed since the early 1910s). Nonetheless, it also suggests that the difference was not constant throughout the fourteen years after its adoption; the homicide rate increased rapidly during the early years of Constitutional Prohibition, and slowly started to fall back to pre-Prohibition levels around 1926.

Observed arrests for drunkenness are the equilibrium outcome of alcohol demand, alcohol supply, and intensity of arrest enforcement. Their evolution captures changes in all of these components. Figure 2 presents the population-weighted average per-capita drunkenness arrest rate for the 255 U.S. cities for which this variable is available throughout the whole 1911-1929 period. Its sharp fall started well before Constitutional Prohibition was adopted. It fell to around 39% of its initial level (from around 18 to only 7 arrests per $1,000$) in just a few years. On the other hand, it was precisely in 1920, the year when the 18th Amendment entered into force, that drunkenness arrests started bouncing back at an even faster rate. They finally converged to around 83% of their average initial level, at a time when federal Prohibition was still in place. The breaks in both the homicide rate and the drunkenness arrest rate series do not appear to match the introduction of Constitutional Prohibition. This suggests differential short-run and long-run effects of Prohibition, and the relevance of state-level Prohibition, which, as mentioned in section 2, occured staggeredly across states during the first two decades of the century.

In fact, throughout the 1910s arrests have a sharp fall in every city, but at different points in time across cities in different states. The fall appears to be highly correlated with the timing of adoption of state-level Prohibition. Figure 3 presents the dates of adoption of state-level Prohibition. It shows how the Prohibitionist wave moved across the United States during the 1910s, up until the introduction of nationwide Prohibition with the 18th Amendment [16].

In the context of alcohol Prohibition, time under the policy is a convenient reduced-form way to look at its time-varying effects for several reasons. First, because of the alcohol supply dynamics; after Prohibition is adopted, the legal market for alcohol is closed on impact. This implies a large negative shock on the availability of liquor. The black market requires time to develop smuggling networks and establish hidden production facilities. Moreover, because crime is a necessary input into the production and trade of any illegal commodity, costly and time-consuming investments are also necessary for the development of criminal organizations supporting the illegal market. Finally, law enforcement was a key channel through which Prohibition had an impact on the development of criminality, and equilibrium law enforcement depended on the community's beliefs about the policy. The evolution of these beliefs over time was also a dynamic force shaping the time-varying effects of Prohibition as a legal standard. Thus, Prohibition is likely to have varying short-run and long-run effects. To obtain an estimate of the overall effects of Prohibition, a comparison of cities which have experienced similar lengths of time under the policy is needed.

To take a first look at short-run and long-run effects of Prohibition, I start by estimating fixed-effects models of the form

$$y_{ct} = \alpha_c + \beta_t + \sum_{\tau=1}^{k} \delta_\tau D_{c\tau} + \boldsymbol{\gamma}' \boldsymbol{X}_{ct} + \varepsilon_{ct} \tag{1}$$

where c indexes cities and t indexes years. y_{ct} can be either the homicide rate, the drunkenness arrest rate, or police expenditure, for which I look at two alternative measures: Police expenditure as a share of total city public expenditure, and per capita police expenditure. The α_c are city-specific effects, the β_t are year-effects, and the $D_{c\tau}$ are indicator variables for each cumulative number of years under Prohibition[17]. The vector $\boldsymbol{X}_{ct}$ includes a constant, the log of population to capture any scale effects, and time-varying effects for border and state-capital indicators. The focus of Equation (1) is in the estimates of δ_τ, the time-varying effects of Prohibition. Since this model looks only at within-city variation over time, the δ_τ can be interpreted as the average-across-cities difference in y_{ct} relative to the city average, when a city has been under Prohibition for τ years. Standard errors reported are robust to arbitrary heteroskedasticity and clustered at the city level to adjust for arbitrary within-city correlation over time. Because of the strong trend in the police expenditure data, I also ran "random trend" models for some specifications for this outcome variable, allowing for city-specific linear trends.

I present regressions for two alternative samples, labeled as B and C[18]. B is a balanced sample

[16] In figure 3, Kansas, Maine and North Dakota are not shown because these three states were already under Prohibition since the late 19th century. Kansas adopted Prohibition in 1880, Maine in 1884, and North Dakota in 1889 (at the same time it acquired statehood). Kansas and Maine had already been under Statewide Prohibition in the mid-1800s during the first Prohibitionist wave.

[17] In the sample τ runs up to 55, given that Kansas was under Prohibition since 1880. Because only very few cities experienced Prohibition for more than eighteen years, I restrict k to be 19, and leave observations with more than nineteen years under Prohibition as part of the omitted category.

[18] I call A the complete sample including all observations for which data is available, and also ran regression on it which I omit from the paper. Thoughout, no significant differences arise from results using either sample.

of the 66 cities for which complete data is available for the whole period 1911-1936, which will be the sample used for the stuctural estimation in sections 5 and 6. Sample C is an unbalanced panel excluding cities for which there are less than ten years of data for drunkenness arrests or police expenditure, or less than eight years of homicide rate data. Thus, $B \subset C \subset A$. The complete regression results can be found in Appendix 4. For brevity and ease of illustration, the left panel in figure 4 graphs the estimated δ_τ's of the baseline specification with no year effects. It nicely shows how the homicide rate is relatively unresponsive for the first few years after a city has been under Prohibition, and then trends upwards until around the 10th year under Prohibition. The homicide rate then starts slowly falling back to a level similar to the pre-Prohibition average. The set of cities experiencing lengthier periods under Prohibition shrinks over time, so late δ_τ's are less precisely estimated. At its peak, cities were on average experiencing 3.1 homicides per $100,000$ more than before Prohibition was introduced (s.e.$=2.7$).

Analogous regression results for drunkenness arrests provede a complementary picture. The estimated δ_τ's are presented in the right panel of figure 4. The figure illustrates the dramatic fall in drunkenness arrests during the first two years after a city was under Prohibition. This is the expected outcome of prohibiting the liquor trade, due to the impact closing of most of the supply sources of alcohol which, during this period, were to a large extent domestic. The reduction in the supply of alcohol is likely to be underestimated in figure 4, given that law enforcement does not show a fall during early Prohibition years, relative to years without Prohibition. During the second year under Prohibition, drunkenness arrests attain a minimum. The estimated coefficient for δ_2 is -9.73 (s.e.$=1.4$), which implies that at its lowest point, the alcohol supply would have contracted 50% ($=12.7/19$) in the absence of changes in law enforcement or demand. The figure also illustrates the steady recovery of the alcohol market, if we are willing to assume that arrest intensity did not change significantly throughout Prohibition. Approximately fifteen years into Prohibition, drunkenness arrests are indistinguishable from Pre-Prohibition levels[19].

The patterns in panel A of figure 4 are consistent with the idea that legal Prohibition immediately had a large effect on the supply of alcohol. When looking at crime, it had a much smaller short-run impact, likely due to the slow development of alternative (illegal) sources of alcohol and their associated crime networks. On the other hand, the figure does not support the claim of Prohibitionists of the time, who claimed Prohibition would reduce criminality and the social disruptions associated with liquor consumption and the saloon; despite the large contraction of the alcohol market during the early prohibitionist years, a time when criminal organizations were still not developed, the homicide rate remained relatively steady.

Finally, panel B in figure 4 presents the estimates of the δ_τ's both for the police share and for the per capita police expenditure. Both measures of law enforcement increase steadily until around ten to twelve years into Prohibition, only to subsequently fall back at a mildly faster pace. The pattern follows the one of the homicide rate; both variables appear to increase during the first years of Prohibition, and to start falling at relatively similar times. Below I will argue that the rise and

[19]The identification assumption here is that the introduction of Prohibition did not also induce changes in individual's preferences over alcohol consumption. As an effort to check how reasonable this assumption is, Appendix 4 presents some evidence exploiting variation in the availability of neighboring alcohol markets. The evidence there is consistent with no changes in demand after the introduction of Prohibition.

fall patterns in police enforcement and crime can be understood as the equilibrium outcomes of a dynamic learning process about the effects of Prohibition, and its interaction with the distribution of moral preferences and the dynamics of the illegal alcohol market and its associated criminal networks[20].

4.2 Preferences and Moral heterogeneity

Communities with varying preferences over the legal standard were likely to collectively respond differently to the introduction of Prohibition. The trends in figure 4 are likely to be averaging out heterogeneous responses across cities with different moral profiles, and beliefs and thus, with differing willingness to enforce the policy. On the one hand, drier constituencies should be willing to enforce more because for the decisive voter, her marginal valuation of reducing her community's alcohol consumption was larger, and because she was likely to be more optimistic about the effects of law enforcement under Prohibition. On the other hand, drier communities were likely to face smaller potential alcohol markets, and hence less crime increases due to Prohibition. Thus, holding moral views constant, the decisive voter in cities with larger drinking populations faced an incentive to increase law enforcement, relative to cities with smaller alcohol markets. If moral views were relatively fixed, changes in equilibrium law enforcement should be due to belief updating about the effects of Prohibition.

The empirical analysis below is based upon comparing changes in outcomes over time, in cities having different distributions of moral tastes, exploiting both the timing of adoption of Prohibition laws and the variation in community preferences. I follow the historical literature, and use the variation in the religious ascription distribution and in the ethnicity and age distribution of the population, as the main observable characteristics correlated with moral views about alcohol Prohibition, and prior beliefs about its effects. I construct a straightforward proxy for the "wet share" in the population, μ_{ct}, as the sum of the fractions of the population in any of the religions considered as "wet" in the literature, the share of non-native white individuals, and the share of the population in the 15-44 years range[21]. There is fairly widespread consensus that Baptist, Evangelical, Methodist, Mormon, and Presbyterian religious ascriptions were more favorable to Prohibition, while Catholic, Orthodox, Jewish, and Lutheran communities had much more positive views about alcohol consumption. On the other hand, while native whites, especially native white women, were strongly prohibitionist, foreign whites (Irish, Italians, Germans, Polish, Scandinavians) and blacks were more liberal about alcohol consumption. Finally, it is likely that younger populations also had more liberal views about

[20]Evidence that Prohibition enforcement was weakened after an "experimentation" also comes from the repeal of enforcement laws in several states during the 1920s, as mentioned in section 2 when discussing the controversy over the repeal of New York's enforcement law. States under Prohibition before the adoption of the 18th Amendment had their own alcohol enforcement legislation, which was in many cases strengthened or harmonized with federal legislation after Congress passed the Volstead Act. All other states, with the exception of Maryland, adopted state-level enforcement legislation right after the passage of the Volstead Act, thus complying with the shared-enforcement responsibilities established by the 18th Amendment. Throughout the 1920s several states decided to repeal their state-enforcement laws, effectively leaving the federal government alone in the enforcement of Prohibition. The state of New York took the lead by repealing its enforcement law in 1923, very much against the will of the Federal government and of a large share of upstate voters. It was followed by Montana in 1925, Nevada and Wisconsin in 1928, Massachussetts in 1930, and Arizona, California, Colorado, Louisiana, Michigan, North Dakota, New Jersey, Oregon, and Washington in 1931.

[21]I normalize this variable dividing by 3, the total measure of the religious, ethnicity, and age distributions.

liquor (See for example Foster (2002); Sinclair (n.d.); Szymansky (2003); Blocker (1989); Asbury (1950)). Thus I define "wetness" as:

$$\mu_{ct} = \frac{1}{3}(1 - \%Baptist_{ct} - \%Evangelical_{ct} - \%Methodist_{ct} - \%Mormon_{ct} - \%Presbyterian_{ct})$$

$$+ \frac{1}{3}(1 - \%NativeWhite_{ct}) + \frac{1}{3}(\%PopulationAges15 - 44_{ct}) \tag{2}$$

In 1911, its mean is 0.49, with a standard deviation of 0.085. Similar to the empirical strategy in Equation (1), I regress each of the outcome variables y_{ct} on the years-under-Prohibition indicators, and their interaction with the initial value of the "wetness" measure [22]. As a benchmark for comparision, I ran analogous regressions using only the Constitutional Prohibition indicator, just as in the models in section 4.1. The models I estimate take the form:

$$y_{ct} = \alpha_c + \beta_t + \sum_{\tau=1}^{k} \delta_\tau D_{c\tau} + \sum_{\tau=1}^{k} \phi_\tau D_{c\tau}\overline{\mu}_c + \boldsymbol{\gamma}'\boldsymbol{X}_{ct} + \varepsilon_{ct} \tag{3}$$

Interest lies in the differential evolution of outcomes over time under Prohibition, captured by the estimates of the ϕ_τ's, which measure how the several outcome variables changed differentially over the years under Prohibition, between cities with varying "moral" distributions (relative to a city with zero "wet" population). For ease of exposition, panel A in figure 5 graphs the estimated ϕ_τ 's for the specifications using sample C. (See table A4-2, column (4), in Appendix 4). Estimates are very similar in magnitude for the alternative samples. The figure shows an increasing differential gap in the homicide rate during the first years under Prohibition, which subsequently closes over time, for cities with relatively "wetter" constituencies. This happened especially during the years in which the homicide rate was high. Because the differential increases in crime followed the same time pattern of overall crime during Prohibition, this suggests that a large fraction of the increase in criminality occured in cities with wetter constituencies. Differential changes in the drunkenness arrest rate, which can be seen in the right panel of figure 5, appear to be small and significantly different from zero only in a few of the years under Prohibition when the alcohol supply was likely experiencing its fastest recovery.

Panel B plots the estimated ϕ_τ's for the police share and per capita police equations. Both show a similar pattern: cities with "wetter" constituencies increased police expenditure differentially less during early Prohibition years, but this gap closes over time, and for later Prohibition years, wetter cities have differentially higher spending in police. The relatively tighter law enforcement in drier cities during the early Prohibition years is consistent with their constituencies having relatively optimistic beliefs about the its effects, making them more willing to repress the alcohol market, and expecting little response of crime. But criminality was increasing relatively more in wetter cities, and

[22] I take 1911 as the baseline value for μ_{ct}. For cities without religious distribution data before that year, I use the earliest year available (1916 in most cases). As a robustness check I ran identical regressions using the 1911 data on the somewhat reduced sample of cities without data before 1916, and results varied only marginally (available upon request).

their alcohol markets were bouncing back faster. This suggests that criminality was very sensitive to the size of the potential alcohol market, requiring higher levels of crime enforcement in wetter cities, despite their preferences for a more lenient enforcement of the Prohibition laws. Indeed, panel B in figure 5 shows that changes in police expenditure were differentially higher in wetter cities during the later years under Prohibition. These were years in which cities were, overall, reducing police expenditure, so the figure implies that wet cities were unable to reduce law enforcement as fast[23].

These patterns suggest that the alignment between the legal standard and community preferences played a major role in determining law enforcement outcomes. In cities where the median individual disfavored alcohol consumption and the alcohol market was small, there was little potential for crime to arise after the introduction of Prohibition; Prohibition enforcement could be tightened without concomitantly high crime increases. If individuals learn about the effects of Prohibition by observing crime outcomes, these communities should not alter their preferences too much over time. In contrast, communities where Prohibition was in stark contrast to average moral preferences over alcohol faced a much more demanding problem. In those cities, the alcohol market was large, so the potential for Prohibition-related criminality was much higher. The median citizen in this community should be unwilling to enforce Prohibition tightly, not only because she was likely to enjoy alcohol consumption and was morally liberal about others' alcohol consumption, but also because she was less optimistic about the response of criminality to Prohibition enforcement. This is what the early behavior of police expenditure suggests in figures 4 and 5, and is consistent with the repeal of state-level enforcement legislation.

If tightening Prohibition enforcement drove illegal producers towards a more intensive use of violence, why did police enforcement fall more slowly in wet cities in the later years under Prohibition, if these were the ones most unwilling to enforce it? I suggest the answer is the impossibility to separate overall crime enforcement and the enforcement of restrictions over a specific market, when the legal standard prescribes full Prohibition. The prohibited market itself becomes a major source of criminality, so that combatting crime also indirectly tightens the alcohol market. Under Prohibition, the ability to specifically target crime without restricting the alcohol market was limited, especially for policing activities. Thus, Prohibition in wet cities not only had adverse effects over crime, but also was costly because for a given level of police expenditure, it would lead to a larger response of crime relative to a city with a smaller alcohol market. This predicts larger shifts in preferences over Prohibition in these communities.

The timing of adoption of State-level Prohibition could be correlated with unobservables at the city level, which themselves would be causing the observed trends. However, this is unlikely because such trends should also have a non-linear effect over time. Moreover, I am looking at the effects of Prohibition on a sample of U.S. cities, which did not directly choose a Prohibitive legal standard, but rather saw it imposed upon them by state and federal decisions, making it less likely that the timing of adoption of Prohibition is correlated with city-specific unobservables. Nevertheless, other variation in previous legislation, in particular other alcohol-related laws, and women's suffrage, appear as potential correlates of the introduction of Prohibition. In Appendix 4 I look at variation in

[23] As a placebo test for the results on police expenditure, I ran analogous models using the expenditure in fire. I do not include the results here to save space, but no discernible differences appear between cities with different moral profiles.

the availability of neighboring alcohol supply sources, in pre-Prohibition state-level alcohol-related legislation, and in women's suffrage legislation. The evidence does not suggest that alternative legislation was driving the patterns described above.

4.3 Public Opinion

To look at changes in political support for Prohibition during this period I exploit alcohol-related referenda election returns, available at the county level for most of the U.S. states, taking place in different years during the 1910s-1930s. I focused on finding for each state, electoral returns on a liquor referendum taking place prior to the introduction of Prohibition in the State (the pre-Prohibition period), and for a year in the later Prohibition period or after the repeal of federal Prohibition (the post-Prohibition period). Because most of the information is available at the county level, here I present results for both a county panel and a city panel, assigning the county vote to the city(ies) in the county[24]. A comparison of the distribution of wet vote shares prior to and after Prohibition reveals the dramatic shift in public opinion. Figure 6 presents the histograms of county wet vote shares in both periods. In the pre-Prohibition referenda, the 75th percentile of the distribution of wet vote shares is 0.5. Thus, in three quarters of the counties some type of Prohibition had majoritarian support. In the post period, only 35% of counties had majorities favoring Prohibition. On the other hand, the comparison of both histograms suggests a spreadout in the distribution of public opinion regarding the policy.

Figure 7 shows differential patterns of opinion shift between communities with varying moral profiles. There was strong convergence of public opinion against Prohibition, but it was restricted to communities that were morally more favorable to alcohol in the first place. The figure breaks the sample of counties between those with a value of my "moral wetness" measure, μ, below (left figure) and above (right figure) the median of 0.355 [25], and plots the pre-Prohibition and the post-prohibition wet vote shares in the horizontal and vertical axes respectively, together with a 45 degree line. Almost all counties above the median had a public opinion shift against Prohibition, while in the set of below median counties, a considerable fraction even observed shifts towards Prohibition. Moreover, among the latter group of counties there is no evidence of "convergence of opinion", since pre-Prohibition vote shares are a very good predictor of post-prohibition ones. In contrast, among above-median counties the shift against Prohibition was on average much larger in counties initially more favorable to Prohibition. The shift in public opinion was concentrated in the upper part of the distribution of moral preferences. Analogous figures for the city sample reveal the same patterns.

More formally, I estimate fixed-effects regressions for both the county and the city samples, with two periods, $t \in \{0, 1\}$. $t = 0$ is the pre-Prohibition period, and $t = 1$ is the post-Prohibition period, for a year in which there was a liquor-related referendum. The models I estimate take the basic form

$$w_{ct} = \alpha_c + \beta t + \delta\mu_{ct} + \phi\mu_{c0}t + \boldsymbol{\gamma}'\boldsymbol{X}_{ct} + \varepsilon_{ct} \quad (4)$$

[24]Except for cities in Massachusetts and Connecticut, for which city-level data is available.

[25]I computed μ for each county directly from equation 2 for county-level data. I used the 1916 and 1926 Census of Religions for the religious ascriptions distribution. For the age and ethnicity distributions I used the 1920 and 1930 Population Censuses because the county-level age distribution from the 1910 census is unavailable.

where w_{ct} is the wet vote share. In this model the interaction term for the post period uses the initial period's wetness, given that it is based on baseline moral preferences that law enforcement and its equilibrium effects are endogenously determined. $\boldsymbol{X}_{ct}$ is a vector of time-varying controls, including the log of population (1910 data for $t = 0$ and 1930 for $t = 1$), the urban share of the county (or of the county's city), the number of dry laws in place, the year in which the referendum took place, and indicator variables for the type of referendum (a Prohibition law, a constitutional convention election or a constitutional amendment (omitted category)). The estimate of ϕ should capture the differential increase in the wet vote share in wetter communities.

Table 3 presents the main results. Columns (1) – (5) present results for the complete sample of counties. For comparative purposes, columns (6) – (10) present estimates for analogous models but restricting the sample to counties with a population larger than 30,000. Finally columns (11) – (15) present results for the sample of cities. Columns (1), (6), and (11) first simply regress the wet vote share on a post-Prohibition period indicator. The estimated coefficient in column (1) implies that the average county experienced a 13 percentage points larger wet vote share after Prohibition (s.e.= 0.004). Column (2) then presents estimates of the main specification in equation (4) without additional controls. Column (3) controls for the log of population and the urban share, the year in which the referendum took place, and indicators for the type of referendum. Both the type of referendum in consideration and the year in which it took place are likely to be endogenous to the vote share, given that the timing and kind of referendum were likely to depend on the trends of public support for Prohibition; for example, a proposal for a constitutional amendment was likely to take place in states where public opinion favoring Prohibition was believed to be widespread. Thus, I do not stress the results of the models in columns (3), (8), and (13); nonetheless, estimates are very similar to those excluding these variables. Column (4) includes state-cross-post Prohibition interactions, and finally column (5) accounts for the potential selection problem arising from the fact that a subset of wet states never held pre-Prohibition liquor referenda, by controlling for the inverse Mills ratio of the estimates of a Probit selection equation for holding a referendum (See Appendix 4). In all regressions I run a completely balanced panel. The estimates of the selection equation are shown in panel B. If anything, the size of ϕ, the estimated differential effect of having a larger wet constituency, increases when accounting for selection.

The estimate of ϕ from column (5) implies that a county with a one standard deviation higher μ_{c0} would differentially increase its wet vote share by 6 percentage points ($0.062 = 0.48 \times 0.13$). The interaccion terms are very precisely estimated across specifications, and the regression results suggest that most of the increase in support for anti-Prohibitionism occured through the differentially larger growth in wet support of morally wet communities. The magnitude and significance of the estimates for the city sample are very close to those of the county sample, as can be seen in columns (11) – (15). Estimates for the restricted sample of more populous counties are even larger in magnitute, and imply that the result is not driven by a comparison of extremely dry versus extremely wet communities.

Given that wetter communities were initially less in favor of Prohibition, there was less room for an increase in anti-Prohibitionist sentiment. Nevertheless, the referenda electoral results suggest these communities did experience larger public opinion shifts. One possibility is that learning in dry communities was slower because of large differences in prior beliefs about the effects of the policy, coupled with uninformative local law enforcement decisions. It could also be that all communities were learning at similar speed, but that differences in moral views were so large that for the driest of communities indirect preferences over the policy were very inelastic to changes in beliefs. Finally, it is also possible that very dry communities did in fact benefit from Prohibition. Although this seems at odds with the reduced-form results on crime presented above, the experience of rural and very dry counties might have been very different, given that local preferences were much more aligned to a prohibitionist legal standard.

Overall the reduced-form results show that the introduction of Prohibition had heterogeneous effects across cities varying in their moral preferences over the policy, and directly point towards a set of elements that a comprehensive theory of endogenous law enforcement in the context of U.S. Prohibition should incorporate. First, that responses varied over time, and that restricting attention to Nationwide Prohibition is insufficient to understand the trends in the different outcomes I focused on; the passage of state legislation and enforcement laws, together with local law enforcement decisions, appear as first order. Second, that learning about the effects of the policy is likely to have driven not only the evident changes in public opinion but also the equilibrium law enforcement choices during Prohibition years. Third, that the dynamics of the alcohol market were important for the evolution of criminality during Prohibition. Finally, that variation in the potential alcohol markets across cities implied differential constraints on the extent to which communties could vary law enforcement.

5 A Statistical Model of Prohibition, Learning, and Endogenous Law Enforcement

In this section I develop a simple political economy model of Prohibition enforcement and learning. It incorporates the central interactions at the heart of the dynamics of criminality and public opinion during Prohibition, based on the discussion above. It provides enough structure to be directly estimated. Importantly, it is an equilibrium dynamic model where equilibrium outcomes are the result of the optimal choices of agents, and where learning is rational. Prohibition altered the Data Generating Process (DGP) of several economic outcomes through two main channels. First, in the absence of Prohibition there is no direct link between law enforcement and criminality; this link arises through the enforcement of dry legislation when Prohibition is adopted. Second, differences in beliefs and uncertainty about the effects of Prohibition created a new dynamic channel affecting law enforcement choices at the local level, because many communities were experimenting a new legal standard with unknown consequences at the time of its adoption. In the model, the interaction between moral preferences and beliefs determines the political-equilibrium choices of law enforcement which, by affecting crime, determines the endogenous evolution of learning about the effects of the policy. The evolution of beliefs subsequenty shift optimal law enforcement choices and public opinion over Prohibition. This requires that individuals know the mapping from indirect preferences to law enforcement choices (the political process), and have beliefs about the mapping from law enforcement to expected outcomes.

5.1 Environment and Preferences

Consider a society made up of a large number of small communities $c = 1, 2, ...$, in discrete time. Community c is populated by a continuum measure 1 of adult citizens indexed by i. Each period $t = 0, 1, 2, ...$, every citizen makes a private decision about alcohol consumption, and through majority voting, collectively decides how to distribute a fixed public budget among public goods. Each adult lives for one period, and has a child[26].

In addition, society as a whole (out of which the individual community is small) can decide a legal standard over the alcohol market for the community, either to be under Prohibition ($P_t = 1$) or not under Prohibition ($P_t = 0$). In the latter regime, alcohol is freely traded (though possibly with some regulation), whereas in the former, an illegal alcohol market is the only source of liquor. Under no Prohibition the alcohol market is perfectly competitive, while under Prohibition, the black market is monopolistic. When Prohibition is in place, the community collectively decides the extent of enforcement of the law. Finally, $P_0 = 0$, so that society's initial legal standard is liberal.

Citizens are heterogeneous in several private and common-values dimensions (Arrow (1963)). In regard to private values, each adult citizen is either dry D_t or wet W_t, and I denote $\mu_t = |W_t|$ as the share of wet adult citizens. The two groups differ in their preferences over individual alcohol consumption h. For simplicity, dry individuals do not derive any utility from their own consumption of alcohol, while wet adult individuals do enjoy consuming a unit of alcohol every period ($h \in \{0, 1\}$). This type is not inherited from parent to child, but during every period the share of wet individuals is a random variable drawn from a beta distribution (See Coate and Conlin (2004) or Degan and Merlo (2009) for a modeling choice in the same spirit):

$$f_\mu(\mu; a, b) = \frac{\mu^{a-1}(1-\mu)^{b-1}}{\int v^{a-1}(1-v)^{b-1} dv}, \quad a, b > 0 \tag{5}$$

Individuals know the parameters of the distribution, but do not observe the draw directly (they do not observe the type of their fellow citizens). Each individual is also characterized by a "moral view" z^i, which is a measure of the marginal disutility she gets from her community-wide alcohol consumption. I will call z^i her moral view, and will assume it is inherited from parent to child.

On the other hand, individuals in the community have common values about consumption of a public good G, and crime, but there is heterogeneity in prior beliefs (beliefs of the cohorts living prior to and during the first period under Prohibition) about how the introduction of Prohibition might impact crime within the community. Thus, conflicting views over Prohibition arise not only from differences in individual moral stands (tastes), but also from informational differences (or differences in the way prior information was interpreted). Nevertheless, these are correlated in the population to allow individuals with more radical views against alcohol consumption (by others) to be more optimistic about the response of criminality to Prohibition.

Specifically, the information structure, which will imply parsimonious learning dynamics, is as follows. Individual i's moral view (distaste for her community's aggregate alcohol consumption) is

[26] Throughout this section I drop the community indices c, since no confusion arises. In section 6 I specify which parameters are city-specific for estimation purposes.

$z^i = z + \zeta^i$, where z is her community's average moral view, and ζ^i is her individual-specific moral shock. On the other hand, her prior beliefs (about the elasticity of crime to the enforcement of Prohibition, as will be explained below) are $\theta_0^i = B + \xi^i$, where B can be thought of as the common component of prior beliefs (which possibly includes a bias), and ξ^i is an individual-specific bias. (ζ^i, ξ^i) is drawn from a joint-normal distribution

$$\begin{pmatrix} \zeta^i \\ \xi^i \end{pmatrix} \sim N\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma_\zeta^2 & \rho\sigma_\zeta\sigma_\xi \\ \rho\sigma_\zeta\sigma_\xi & \sigma_\xi^2 \end{pmatrix} \right) \tag{6}$$

Here moral views are understood as the set of beliefs about the world, which an individual takes as true. This is, to which she assigns a degenerate prior probability of 1, and are thus not subject to updating with the arrival of new information. On the other hand, all other beliefs can evolve through rational updating as the individual receives new information. In the context of Prohibition, it is natural to think of crime as the source of information about θ. Observe that if $\rho < 0$, individuals who have stronger moral views against alcohol will be on average more optimistic about the response of crime to the introduction of Prohibition. For simplicity, both wet and dry individuals get their (ζ^i, ξ^i) drawn from the same distribution.

The expected utility of a citizen is given by

$$E_t U^i(h_t^i, A_t, G_t, q_t | P_t) = E\left[1_{\{i \in W_t\}} h_t^i - z^i A_t + V(G_t) - q_t\right] \tag{7}$$

where A_t is the aggregate alcohol consumed in his community, q_t is the crime rate, $G_t \in [0, 1]$ is the share of the public budget allocated to public goods other than policing, and E is the expectations operator conditional on all the information available to individual i. The term $-z^i A_t$ represents the "moral externality". Finally, $V(G) = exp(G)$. Notice that from the point of view of individuals the optimization problem is static, since they only live for one period.

5.2 The Alcohol Market

Imagine a very simple alcohol market, where the price of consuming a unit of liquor is normalized to zero under no Prohibition, but individuals must engage in a costly search. The probability of successful search is a decreasing function of the level of Prohibition enforcement chosen by the community[27]. As communities decide to tighten enforcement of dry laws, the availability of alcohol is diminished. Specifically I allow this probability to take the form $Pr(h_t^i = 1 | P_t = 0) = exp(-e_t)$ where $e_t \geq 0$ is the level of dry law enforcement.

The introduction of Prohibition, on the other hand, makes legal alcohol unavailable (and increases market power since black markets are likely to be captured by a small set of criminal organizations). The search for alcohol becomes costlier, and I will allow the probability of a successful search to also become a function of the amount of time the community has been under Prohibition, τ_t, to flexibly capture the possibility that the illegal market adjusts over time. After Prohibition is adopted, the

[27] Recall that under no Prohinition dry laws were in place. These restricted the availability of liquor by regulating the alcohol market along different dimensions.

legal market for alcohol is closed on impact, which by itself has an effect on the quantities traded. The supply response from illegal producers does not occur immediately because it takes time to build up a black market, and the development of crime networks associated with the illegal activity also requires costly and staggered investments. Thus, the probability of successful search under Prohibition is given by$Pr(h_t^i = 1 | P_t = 1) = k(\tau_t)exp(-e_t)$, where

$$k(\tau_t) = 1 - \lambda \tau_t exp(-\kappa \tau_t) \tag{8}$$

with $\kappa, \lambda > 0$[28]. It follows that aggregate alcohol consumption is

$$A_t(e_t) = \int_{i \in W_t} 1k(\tau_t)exp(-e_t)di = \mu_t k(\tau_t)exp(-e_t) \tag{9}$$

so that during the τth year under Prohibition, holding law enforcement constant, the alcohol market is a fraction $k(\tau_t)$ of what it would be under no Prohibition. This highlights why individuals with moral views opposed to alcohol might want to choose high levels of law enforcement. By reducing the equilibrium consumption of alcohol, their moral externality is directly reduced. The fact that after an initial fall $k(\tau_t)$ rises as time under Prohibition increases, implies that over time, higher levels of law enforcement are required to maintain a given size of the illegal alcohol market.

5.3 Crime, Prohibition, and Law Enforcement

I allow crime to be related to alcohol consumption by assuming that baseline crime is proportional to the size of the alcohol market. Formally,

$$q_t^N = \Theta_S + A(e_t) = \Theta_S + k(\tau_t)\mu_t exp(-e_t) + \varepsilon_t \tag{10}$$

where $\varepsilon_t \sim N(0, \sigma_q^2)$ is an iid normally distributed shock. Because the homicide rate levels vary significantly across states but are relatively similar between cities in the same state, I allow for a state-specific parameter Θ_S. Central to the understanding of the variation in criminality across the United States during Prohibition is the fact that different communities were structurally different in how the ban on the alcohol trade would affect criminality, and there was disagreement about this issue. I will assume the following relationship between law enforcement and Prohibition-related crime:

$$q_t^P = \theta[A_t(e_t = 0) - A_t(e_t)] = \theta k(\tau_t)\mu_t[1 - exp(-e_t)] \tag{11}$$

[28] I introduce two parameters for $k(\tau_t)$ to be flexible enough to separately capture the initial fall in the alcohol market once Prohibition is enacted (λ), and the speed at which the alcohol market bounces back (κ), and will restrict them to be constant across cities in the empirical analysis below. Note that for no-Prohibition years, $k(\tau_t) = k(0) = 1$. A graph of $k(\tau_t)$ is presented in the first panel of figure 11 for $\kappa = 0.26$ and $\lambda = 0.25$ (the MLE estimates). This curve has its unique minimum at $\tau_t = 1/\kappa$, so that κ is also the inverse of the time at which the alcohol market reaches its mimimun size. I also impose the condition $\kappa exp(1) > \lambda$, which is necessary and sufficient for $k(\tau_t)$ to be everywhere positve. A comparison of figures 2 and 11 illustrates why the functional form choice in 8 is likely to be appropriate.

Equations (10) and (11) capture the two main channels from the alcohol markt to crime. Alcohol consumption can cause crime by altering the behavior of consumers, and by giving incentives for the development of crime networks when it is prohibited [29]. Total crime is $q_t = q_t^N + P_t q_t^P$. In equation (11), θ is the true state, a city-specific shifter of crime to the size of the alcohol market under Prohibition. Formally, this implies a structural change in the Data Generating Process for crime when Prohibition is introduced. For $\theta > 0$, it measures the extent to which crime increases as the alcohol market is tightened through law enforcement, relative to the size of the market at zero law enforcement. Observe that $q_t^P = 0$ if $e_t = 0$, or under no Prohibition. Also, as $e_t \to \infty$, Prohibition-related crime $q_t^P \to \theta_c k(\tau_t)\mu_t$. This functional form captures a set of key aspects about the link between criminality and law enforcement under Prohibition. First, sustaining a smaller black market when alcohol is prohibited, translates into more crime. Second, a larger wet share implies a larger potential alcohol market, and hence, more Prohibition-related crime for a given level of law enforcement. Third, the time-variation in crime should be correlated with the time-variation in the alcohol-market dynamics. Fourth, and most importantly, a link between restrictions in the alcohol market and criminality only appears when Prohibition is in place. There is common knowledge up to the uncertainty about the value of θ.

The drunkenness arrest rate is, by definition, the conditional probability of being arrested times the alcohol market size. It is a function of law enforcement, and I will allow the probability of being arrested to take the flexible form $Pr(Arrest|e_t) = \frac{exp(e_t)}{\chi + exp(e_t)}$, with $\chi > 0$[30]. The drunkenness arrest rate is thus:

$$d_t = Pr(Arrest|e_t)A_t(e_t) = \frac{\mu_t k(\tau_t)}{\chi + exp(e_t)} \tag{12}$$

Notice this equation holds both under no Prohibition and under Prohibition, since under no Prohibition public drunkenness was also prosecuted. The equilibrium drunkenness arrest rate is a decreasing function of law enforcement. Equation (12) highlights that variation in the drunkenness arrest rate can come from changes in the size of the alcohol market, (the wet share μ_t and the "secular" dynamics of the alcohol supply under Prohibition $k(\tau_t)$), or from the extent of law enforcement e_t. Moreover, when identifying these two channels separately, the structural estimation will exploit the common variation in drunkenness arrests, crime, and police expenditure due to changes in the size of the alcohol market and in law enforcement.

[29] In a classic Sociology paper, Paul Goldstein discusses the different channels from drug use to violence. The author identifies two sources of criminality in a no Prohibition environment: psychopharmacological and economycally compulsive: In the former, "... some individuals, as a result of short or long term ingestion of specific substances, may become excitable, irrational, and may exhibit violent behavior". In the latter, "...some drug users engage in economically oriented violent crime, e.g., robbery, in order to support costly drug use.... Violence generally results from some factor in the social context in which the economic crime is perpetrated." Then he identifies systemic violence as an added source of crime under Prohibition: "... the aggressive patterns of interaction within the system of drug distribution and use... 1. disputes over territory between rival drug dealers. 2. assaults and homicides committed within dealing hierarchies as a means of enforcing normative codes. 3. robberies of drug dealers and the usually violent retaliation by the dealer or his/her bosses. 4. elimination of informers. 5. punishment for selling adulterated or phony drugs. 6. punishment for failing to pay one's debts. 7. disputes over drugs or drug paraphernalia. 8. robbery violence related to the social ecology of copping areas." Goldstein (1985, pp.146-149)

[30] The choice of this logistic functional form for the conditional probability of being arrested under drunkenness charges is flexible enough to allow any arrest probability at zero law enforcement: $Pr(Arrest|0) = 1/(1+\chi)$, which is a convenient way to interpret χ.

Prohibition enforcement is a function of the amount of police expenditure p_t, and the current legal standard, which includes dry laws, enforcement laws, and Prohibition. I will assume Prohibition enforcement can be expressed as $e_t = \alpha_t p_t$, with $\alpha_t > 1$, which depends on the legal standard in place. The multiplicative form is intended to capture the inherent non-separability between crime and Prohibition enforcement. Observe, nonetheless, that liberalizing the legal standard (by lowering α_t) weakens the link between both, at the cost of reducing the restrictions on the alcohol market. Each community has a unit of public resources to allocate between policing p_t and other public goods G_t, and I assume, for simplicity, they can be exchanged one-for-one. Thus,

$$G_t = 1 - p_t \tag{13}$$

5.4 Learning and the Timing of Events

To make the model suitable for estimation, I make the following assumptions about information, learning, and the timing of events. In the end of period $t-1$, each member of the adult cohort has one child, and outcome variables $(p_{t-1}, q_{t-1}, d_{t-1})$ are realized. Under no Prohibition there is no learning taking place, whereas in a Prohibition year, children observe the vector of outcome variables and update their beliefs about θ according to Bayes' rule. This occurs as follows. First, each child learns her parent's belief θ^i_{t-1}. In the first year under Prohibition ($\tau_{t-1} = 1$), child i knows that $\theta^i_0 = B + \xi^i$ (of course, she does not observe B or ξ^i separately), and knows that $\xi^i \sim N(0, \sigma^2_\xi)$ is the marginal distribution of biases in the population. As a result, child i's prior about θ is given by $\theta^i_{t-1} \sim N(\theta^i_0, \sigma^2_\xi)$.

From equation (12), after the child has observed d_{t-1} and p_{t-1}, she can perfectly back-up the realization of μ_{t-1}. Thus, in the public signal $q_{t-1} = \Theta_S + \mu_{t-1} k(\tau_{t-1}) exp(-\alpha_t p_{t-1}) + \theta k(\tau_{t-1})\mu_{t-1}[1 - exp(-\alpha_t p_{t-1})] + \varepsilon_{t-1}$, the only remaining uncertainty comprises the true value of θ and the distribution of ε_{t-1}. It follows that Bayesian individuals' posteriors about θ will be normally distributed. Normal updating will keep taking place cohort after cohort as long as the community is still under Prohibition. Thus, iteratively using normal updating and exploiting linearity of conditional distributions under normality, cohort t's posterior (or $t+1$'s prior) will be given by

$$\theta^i_t \sim N\left(\frac{\frac{1}{\sigma^2_\xi}}{\frac{1}{\sigma^2_\xi} + \frac{1}{\sigma^2_q}\sum_{s=s_0}^{t-1}\omega^2_s}\theta^i_0 + \frac{\frac{1}{\sigma^2_q}}{\frac{1}{\sigma^2_\xi} + \frac{1}{\sigma^2_q}\sum_{s=s_0}^{t-1}\omega^2_s}\sum_{s=s_0}^{t-1}[q_s - \Theta_S - \mu_s k(\tau_t) exp(-\alpha_t p_s)]\omega_s, \frac{1}{\frac{1}{\sigma^2_\xi} + \frac{1}{\sigma^2_q}\sum_{s=s_0}^{t-1}\omega^2_s} \right) \tag{14}$$

where s_0 is the first year in which community c is under Prohibition, and where $\omega_t \equiv k(\tau_t)\mu_t[1 - exp(-\alpha_t p_t)]$ is a measure of the degree of informativeness of the signal [31]. This posterior will be the

[31] Equation (14) above highlights the convenience of assuming normality for both the prior on θ and the conditional likelihood of the signal which, being an affine information structure, results in a very parsimonious learning process where posterior conditional expectations are linear in the signal sequence, making estimation relatively straightforward. Although this seems to be a very restrictive set of assumptions about the information structure and of the cognitive requirements of the learning process, these features of normal learning are actually fairly robust to alternative specifications. For example, if agents are not fully Bayesians, and are limited to making the best linear predictions based on the signal sequence $\{q_i\}_{s0=0}^{t-1}$, their prediction of the conditional mean will exactly match the posterior mean under normal updating, no matter the true data generating process (See for example, Vives (2010, p. 379)).

relevant measure with respect to which individual i will evaluate her expected utility under different law enforcement policy alternatives.

The stochastic process in (14) is a bounded martingale, and as such, the $\{\theta_t^i\}$ converge almost surely as $t \to \infty$. Moreover, because the true distribution (a mass point of 1 at θ) is absolutely continuous with respect to the prior (which is normal and hence has positive density everywhere), the process will converge to the true θ for any infinite sequence of positive $\{\mu_t, e_t\}_{t=s_0}^{n}$. How rapidly convergence occurs will depend on the amount of law enforcement. As $p_t \to 0$, the signal becomes uninformative because individuals know the data generating process, and hence, realize that at zero enforcement any observed crime rate must not come from Prohibition-related crime. Conversely, for a given observed signal, a higher value of law enforcement reduces the variance of the signal's likelihood, making its informational content much higher. Rational individuals should then put a higher weight on such a signal. Interestingly, this implies that if a community reduces its enforcement levels, it will also reduce the speed at which its members will be able to learn about the true state. Now I define $\overline{\theta}_t^i$ as the posterior mean, and express it more compactly as

$$\overline{\theta}_t^i \equiv \Omega_t \frac{1}{\sigma_\xi^2}\theta_0^i + \Omega_t \frac{1}{\sigma_q^2}\sum_{s=s_0}^{t-1}[q_s - \Theta_S - \mu_s k(\tau_t) exp(-\alpha_t p_s)]\omega_s = \Omega_t \frac{1}{\sigma_\xi^2}\xi^i + \Omega_t \overline{\theta}_t^C \tag{15}$$

where $\Omega_t \equiv \frac{1}{\frac{1}{\sigma_\xi^2}+\frac{1}{\sigma_q^2}\sum_{s=s_0}^{t-1}\omega_s^2}$ is the posterior variance, and the common component of beliefs (shared by all individuals in the community) is $\overline{\theta}_t^C \equiv \frac{1}{\sigma_\xi^2}B + \frac{1}{\sigma_q^2}\sum_{s=s_0}^{t-1}[q_s - \Theta_S - \mu_s k(\tau_t) exp(-\alpha_t p_s)]\omega_s$. The posterior mean belief at any time t is a weighted average of the prior mean and the whole history of crime realizations, weighted according to their relative precisions and by the informativeness of each signal. The degree of informativeness depends, in turn, on the extent of law enforcement originating the signal. Equation (15) shows that individual belief sequences can be analytically decomposed into a common component, shared by all individuals in the community every period given the public nature of the signal, and an individual-specific component, tied to the dynasty-specific bias. Of course, individuals do not separately observe the public and the private components of their beliefs, but the explicit distinction will be convenient. Equation (15) is readily interpretable. When the precision of the distribution of prior biases is low (as measured by $1/\sigma_\xi^2$), Bayesian individuals will disregard the information in their prior and will rely more closely on the observed signal sequence. A lower precision of the signal ($1/\sigma_q^2$) induces a Bayesian individual to put more weight on her prior. Moreover, since individuals know the DGP up to the uncertainty about θ, they optimally use the information on law enforcement to decide how much weight to give to the crime signal.

5.5 Political Equilibrium and the Distribution of Preferences over Law Enforcement

5.5.1 The Problem under no Prohibition

Replacing the probability of successful search, and equations (9), (10), and (13) into (7), indirect preferences under no Prohibition can be obtained. The first order condition implies that The preferred police enforcement of individual i is given by (see Appendix 1),

$$p_t^*(\zeta^i) = \frac{1}{\alpha_t - 1}\left\{ln\alpha_t + ln\left[\frac{a}{a+b}(z+1) + \frac{a}{a+b}\zeta^i - 1_{\{i\in W_t\}}\right] - 1\right\} \tag{16}$$

If the expression inside $ln[]$ is negative, $p_t^*(\zeta^i) = 0$. This expression follows from the fact that μ_t is distributed $\beta(a,b)$, so its mean is given by $\frac{a}{a+b}$, and that the expected alcohol consumption for a wet individual is equal to the probability of successful search. When a community is not under Prohibition, beliefs about θ do not appear in the objective function of its members. The ideal choice of police enforcement simply trades off the reduction in other public goods with the reduction in moral externality from tightening the alcohol market, and the reduction in overall crime. Individuals with higher z^i will prefer higher levels of law enforcement.

Equation (16) illustrates clearly some of the interesting interactions in the context of moral conflict. Wet individuals, who suffer a small moral externality from average alcohol consumption, prefer low levels of policing to reduce the size of the market, but differentially higher the larger is the alcohol market in their community (the larger is $a/(a+b)$). Interestingly, this interaction effect is not present for dry individuals; for them, the marginal disutility of a larger alcohol market induced by a reduction in policing is exactly offset by the marginal disutility of increased criminality brought about by such a reduction in crime enforcement. The effect of tightening the legal standard on the ideal choice of policing, on the other hand, is ambiguous, since it trades off the value of reducing expenditure in police with the complementarity of police enforcement and the legal standard. For large values of α_t though, ideal policing is falling in α_t.

5.5.2 The Problem under Prohibition

Taking a look at the problem under Prohibition by replacing the successful-search probability and equations (9), (10), (11) and (13) into (7), indirect preferences under Prohibition are obtained. From the first order condition, the preferred police enforcement of individual i under Prohibition is given by (See Appendix 1),

$$p_t^*(\zeta^i,\xi^i) = \frac{1}{\alpha_t - 1}\left\{ln\left[\alpha_t k(\tau_t)\right] + ln\left[\frac{a}{a+b}\left(z - \Omega_t\overline{\theta}_t^C + 1\right) + \frac{a}{a+b}(\zeta^i - \Omega_t\frac{1}{\sigma_\xi^2}\xi^i) - 1_{\{i\in W_t\}}\right] - 1\right\} \tag{17}$$

where I have made use of equation (15). Once again, if the expression inside $ln[]$ is negative, $p_t^*(\zeta^i,\xi^i) = 0$ is the preferred police enforcement share. What matters for individual i is his mean belief about θ. Under Prohibition, individuals must now include the increased criminality induced by Prohibition enforcement in their optimal trade-off regarding police expenditure. Equation (17) highlights that the introduction of Prohibition alters individuals' optimal degree of law enforcement, which now becomes a function not only of their wet or dry identity and their dynasty-specific moral shock ζ^i, but also of their dynasty-specific belief bias ξ^i. These are the three sources of unobserved heterogeneity in the model.

The analysis above looked at the indirect preferences of individuals over law enforcement. Nevertheless, law enforcement is a collective decision, which is made through majority voting. Thus, I define an equilibrium of this model as follows:

Definition. *An equilibrium is a sequence of police expenditure shares* $\{p_t^*\}_{t=0}^{\infty}$ *such that for every* t, p_t^* *wins any pairwise vote against any other* p_t' *when all adult citizens vote sincerely given their current beliefs* $F_t^i(\theta)$, *sequences of homicide and drunkenness arrest rates* $\{q_t\}_{t=0}^{\infty}$, $\{d_t\}_{t=0}^{\infty}$ *given by* (11) *and* (12), *and a sequence of belief distributions* $\{F_t^i(\theta)\}_{t=0}^{\infty}$ *for each dynasty* i, *which are updated every period according to Bayes' rule and given by* (14).

To find the equilibrium path, it is necessary to look at the collective decision-making process, which takes the form of simple majority voting. Although there are three sources of heterogeneity regarding preferences over law enforcement across individuals in this model, below I show they can be reduced to one dimension, over which a unique majority-voting equilibrium exists.

Proposition 1. *For any* t, *a given distribution of beliefs* $F_t^i(\theta)$ $\forall i \in [0,1]$, *and a legal standard vector* (α_t, τ_t, P_t), *there is a unique equilibrium level of law enforcement* p_t *given by*

$$p_t^* = \frac{1}{\alpha_t - 1}\left\{ln\left[\alpha_t k(\tau_t)\right] + ln\left[\frac{a}{a+b}\left(z - P_t\Omega_t\overline{\theta}_t^C + 1\right) + (1-P_t)\varrho_N^{med} + P_t\varrho_P^{med}\right] - 1\right\} \tag{18}$$

where ϱ_N^{med} *and* ϱ_P^{med} *are random variables whose densities* $f_{\varrho_N^{med}}(\varrho_N^{med}; a, b, \sigma_{\varrho N})$ *and* $f_{\varrho_P^{med}}(\varrho_P^{med}; a_c, b, \sigma_{\varrho P t})$ *are continuous and positive over the interval* $[-1, 0]$.

Proof. See Appendix 2. □

As the proof of Proposition 1 shows, ϱ_N^i and ϱ_P^i are one-dimensional sufficient statistics capturing the three sources of heterogeneity in individual i's preferences, during no Prohibition and Prohibition periods, respectively. Their conditional distribution across the population is a mixture of two normal densities, weighted by the wet share μ_t. In Appendix 2 I show that the equilibrium level of law enforcement is determined by the median voter's value of ϱ_j^i. Because the wet share is itself a beta-distributed random variable, ϱ_j^{med} is also a random variable whose equilibrium density $f_{\varrho_j^{med}}(\varrho_{Nj}^{med}; a, b, \sigma_{\varrho_j})$ is continuous and takes positive values over the interval $[-1, 0]$. As $\mu_t \to 1$, $\varrho_j^{med} \to -1$, and as $\mu_t \to 0$, $\varrho_j^{med} \to 0$. When all the community is wet, for example, $\mu_t = 1$ so the median in the community corresponds to the median over the distribution of preferences of wet individuals. These are normally distributed with mean and median at -1, given the preference for private alcohol consumption of wets.

5.6 Predictions and Main Assumptions

5.6.1 Predictions

The model makes several predictions about the equilibirum dynamics of law enforcement during Prohibition. First, observe the dynamic trade-off faced by a relatively liberal median voter living in a relatively wet community. Her pessimism about Prohibition-related crime ($\overline{\theta}_t^i > 0$) makes her prefer a low level of law enforcement. This reduces expected crime and increases the likelihood of alcohol consumption. But over time, maintaining a weak law enforcement becomes costlier because overall crime will be rising fast as the alcohol market catches up and its associated crime networks develop over time (as captured by $k(\tau_t)$). This individual is constrained by a lack of independent policy instruments; by maintaining a low level of Prohibition enforcement, she is at the same time reducing overall crime enforcement. Moreover, the trade-off is more demanding the wetter the community, because a median voter in a wet community is more likely to have a liberal stand on Prohibition and be pessimistic about its effects, while facing a larger alcohol market.

On the other hand, the evolution of preferred law enforcement under Prohibition is determined by the difference between moral views and beliefs, both in their common $(z - \Omega_t \overline{\theta}_t^C)$ and individual (ϱ^{med}) components. From the martingale property of the stochastic updating process, $\Omega_t \overline{\theta}_t^C \rightarrow_{a.s.} \theta$. Nevertheless, the informativeness of signals, as measured by ω_t, is increasing in law enforcement. Thus, early law enforcement choices are likely to be low, making early signals uninformative. Moreover, ω_t is also increasing in $k(\tau_t)$, so the relatively small alcohol market of early Prohibition years also reduces the informativeness of signals for a given level of law enforcement. As a result, learning should be slow during the first years under Prohibition, implying that the incentives to increase law enforcement as the alcohol market catches up are likely to dominate the incentives to reduce law enforcement due to changes in beliefs. As the market converges to its pre-Prohibition size (recall $k(\tau_t) \rightarrow 1$ as τ_t increases), and the increased levels of law enforcement increase the precision of the signals, learning will be faster and incentives to shrink law enforcement due to an increasing sequence of beliefs $\overline{\theta}_t^i$ should dominate. After its initial fall, an invert U-shaped pattern of law enforcement should be observed.

Now, notice the presence of the term $k(\tau_t)$ in equation (18). *Ceteris paribus*, law enforcement should fall discretely right after Prohibition is introduced. This is the optimal response to the sharp contraction of the alcohol supply when it is closed on impact. Not only is the potential for crime small because the size of the alcohol market is smaller, but the marginal moral disutility of reducing law enforcement is also low because the alcohol market has sharply contracted, making it optimal to reduce policing. This increases the consumption of other public goods G_t and the private utility of alcohol consumption for wet individuals. Moreover, if prior beliefs for the median individual are such that $\theta_0^i > 0$, her ideal choice of law enforcement would fall even further because she is pessimistic about the response of crime to Prohibition.

In addition to these time-series predictions, equation (18) also makes predictions about the cross-sectional variation in law enforcement and learning. Specifically, variation in average moral views and alcohol market sizes across cities should interact with the evolution of beliefs. From equation (18), individuals in communities with higher average moral disutility (larger z) should be less sensitive to changes in beliefs than individuals in communities where mean moral views are more favorable to alcohol. Thus, morally drier communities should be more reluctant to change law enforcement as learning takes place. This suggests a nuanced inverted U-shape pattern of ideal law enforcement in relatively dry cities. Of course, such a pattern across cities could be alternatively interpreted as arising from behavioral differences in the ability to learn, between individuals with differing moral views. This model can accomodate equilibrium differences in response to learning while still fully assuming rational individuals.

The correlation parameter ρ has interesting implications in the model. A high correlation between individual moral views and prior biases implies that relative to no-Prohibition years, during Prohibition the decisive voter's preferences will be more extremist, so that an amplification in the difference between the equilibrium choices of drier and wetter cities should be observed. Conversely, if this correlation is low, the average draw of ϱ^{med} will be very similar in Prohibition and no Prohibition periods, so that changes in law enforcement should not vary significantly between dry and wet communities when the legal standard is reformed.

5.6.2 Crime

Equations (10) and (11) are intented to capture some key features of the relationship between crime and the alcohol market. In the baseline equation for crime (10), I introduce Θ_S, a scale parameter at the state level, to capture the large differences in the homicide rate levels across states. Following the claims of Prohibitionists, who argued that alcohol consumption was a source of criminality and social disruptions, I also allow it to vary with the size of the alcohol market.

Regarding Prohibition-related crime, a main reason why large cities were forced to maintain high police enforcement levels during Prohibition was their large potential for criminality, if policing were to be weakened. This points to a central conflict that arises in the context of Prohibition. The enforcement of a prohibitionist legal standard creates a non-separability between the objective of enforcing Prohibition and of controlling crime. The instruments for the enforcement of Prohibition, mainly policing and judicial prosecution, are the same used to fight crime at the local level. If the enforcement of Prohibition creates crime, a community that does not favor Prohibition is unable to reduce law enforcement because it cannot be weakened without, at the same time, weakening overall crime enforcement. This is an especially binding constraint in relatively wet communities where criminality is more responsive to falls in crime enforcement, and motivates the functional form in equation (11).

It is frequently argued that crime increases during Prohibition were due to a shift of resources from crime protection to Prohibition enforcement. But this would predict exactly the opposite patterns to those observed in the data. It cannot explain why the steepest increases in crime and law enforcement were observed precisely in the wettest cities in the United States, since it would imply that relatively wet communities, strongly opposed to Prohibition, should have kept their Prohibition enforcement at very low levels and their crime enforcement resources high. This would have avoided a rise in criminality, and would have allowed the black market to operate with relative freedom. On the other hand, if the enforcement of Prohibition cannot be fully separated from overall crime enforcement, then wet communities must have been unable to reduce Prohibition enforcement. Although likely to have a median voter more willing to invest in Prohibition enforcement, relatively dry communities faced smaller alcohol markets. Thus, they faced a lower potential for crime increases if law enforcement were to be weakened. These predictions are consistent with the patterns in the data[32].

[32] This is not to say that a crowding-out of crime enforcement did not take place as communities increased the resources allocated to Prohibition enforcement. Indeed, the widely acknowledged congestion in judicial courts due to Prohibition-related cases is a good example of how it did to some extent shift resources away from overall crime enforcement. The argument here is just that the crowding-out had second order effects relative to the problem arising from the inherent difficulty in separating the enforcement of overall crime and of alcohol Prohibition.

Equation (11) also assumes that crime under Prohibition is a linear function of θ. This is a relatively weak assumption, given that even if it is not linear, equation (11) could be seen as a first order approximation to any other nonlinear structural relationship between q_t^P and the wedge in the alcohol market arising from Prohibition enforcement. Under such an interpretation, the error term would be capturing approximation error. Thus, any misspecification of this relationship should show up in the standard errors of the parameter estimates of equation (11).

5.6.3 Drunkenness Arrests

The functional form specifying the relationship between the drunkenness arrest rate and law enforcement implicitly assumes that throughout the relevant range of law enforcement intensities, the alcohol supply falls at a faster rate than that at which the arrest probability increases. It is adopted for simplicity only, since it makes the derivation of the conditional likelihood more straightforward, by allowing the mapping from unobserved variables to outcomes to be one-to-one for the whole range of outcomes. Moreover, the data suggests this is a reasonable assumption, since variation in law enforcement is only midly correlated with variation in the drunkenness arrest rate, while the timing at which we know the market must have contracted is highly correlated with it across the sample.

5.6.4 Learning

In the model each dynasty gets a specific bias ξ^i, which is analogous to assuming heterogeneous priors in the population. Following Sethi and Yildiz (2009), ξ^i can represent all the information which individual i finds relevant about θ, but is seen as irrelevant for everybody else. The historical literature has emphasized that initial public opinion regarding the effects of Prohibition was extremely optimistic. These biases came from two main sources -some relatively successful experiences of States that underwent Prohibition in the second half of the 19th century, and more importantly, the massive wave of prohibitionist campaigning and lobbying of the ASL and the WCTU in the decades prior to the adoption of nationwide Prohibition (See Asbury (1950); Blocker (1989); Foster (2002); Okrent (2010); Szymansky (2003)). I also assume that both wets and drys get their draw of (ζ^i, ξ^i) from same distribution. This is just a simplifying assumption since, for example, if wet individuals were to get their draw from a mean-shifted distribution, it would be isomorphic to increasing the difference in the marginal utility of private alcohol consumption between wet and dry voters.

Individuals only learn based on local information. This is in opposition to the experimentation literature where learning takes places from neighbors (for example, see Buera et al. (2010)). In the context of Prohibition it is likely that individuals were observing the crime outcomes of other cities. Nonetheless, it is very unlikely that they also observed the local law enforcement choices of other communties. Even if individuals believed that the effects of Prohibition were homogeneous across cities, a signal coming from a city from which law enforcement is not observed is void of informational content. Thus, in this model learning relies on local information exclusively, not only because it is likely that people recognized that Prohibition should have different effects in different communities, but also because learning from signals emerging from unknown law enforcement decisions is not possible without additional information. The endogenous nature of signals in this model justifies that learning should take place based exclusively on local information.

5.6.5 Political Environment

The political equilibrium of this model relies on two main assumptions. First, on simple mayority voting as the collective chioce mechanism. In the context of Prohibition in the United States, bipartisan political competition and a strong involvement of citizens in local politics were prevalent both at the local and federal levels. Indeed, political competition was much weaker in the South during the 1910-1930s, where the Democratic Party had a fairly generalized control of political power. Nevertheless, alcohol Prohibition as a political issue actually increased party competition by making dry voters, who were highly mobilized, involved in politics, and constantly motivated by dry organizations, pivotal [33].

Second, on the absence of any strategic experimentation considerations by voters. Because individuals live for only one period, they simply vote for the level of law enforcement which maximizes their current payoff given their present belief. In a more complex model, one could imagine long-lived or intergenerationally-altruistic voters making strategic voting decisions to induce experimentation in the collective choice of law enforcement level. In the context of Prohibition this is highly unlikely for several reasons. Foremost, local politicians' incentives to experiment with law enforcement were very weak, since adverse criminality outcomes derived from "wrong choices" were likely to hurt their political careers. Indeed, as in any other experimentation setting, experimenting creates positive externalities since learning today benefits not only current, but also future constituencies, and thus, will in general be undersupplied by current constituencies (or politicians). Moreover, in voting environments, incentives to vote for experimentation (in the context of Prohibition higher levels of law enforcement) are weakened by the fact that pivotal voters under the present distribution of beliefs are likely to lose their decisive position after large changes in beliefs induced by experimentation (Strulovici (2010)).

6 Structural Estimation

The equilibrium-political economy model developed in the previous section is characterized by three equilibrium relationships and the dynamic path of beliefs implied by Bayesian updating, which constitute the Data Generating Process (DGP) and can be directly used for estimation (recall that $k(0) = 1$, and c indexes cities):

$$q_{ct} = \Theta_S + k(\tau_{ct})\mu_{ct} exp(-\alpha_{ct} p_{ct}) + P_{ct}\theta_c k(\tau_t)\mu_{ct}\left[1 - exp(-\alpha_{ct} p_{ct})\right] + \varepsilon_{ct} \tag{19}$$

[33] A good example of how competition for the dry vote in the South did increase the competitiveness of local politics was the 1910 Tennessee gubernatorial election. The unwillingness of the incumbent Democratic governor Patterson to enforce the 1909 State Constitutional Amendment introducing Prohibition (after vetoing the Amendment and having his veto overridden by the legislature) alienated a dry fraction of the Democratic party, even after he stepped down for reelection. After more than 30 years in which the Republican party had not occupied Tennessee's gubernatorial office, Republican candidate Ben Hooper won the election on a prohibitionist platform (See Isaac (1965) for a historically detailed account of Prohibition politics in Tennessee).

$$d_{ct} = \frac{\mu_{ct} k(\tau_{ct})}{\chi + exp(\alpha_{ct} p_{ct})} \tag{20}$$

$$p_{ct} = \frac{1}{\alpha_{ct} - 1} \left\{ ln\left[\alpha_{ct} k(\tau_{ct})\right] + ln\left[\frac{a_c}{a_c + b}\left(z_{ct} - P_{ct}\Omega_{ct}\overline{\theta}_{ct}^C + 1\right) + (1 - P_{ct})\varrho_N^{med} + P_{ct}\varrho_P^{med}\right] - 1\right\} \tag{21}$$

$$\overline{\theta}_{ct}^C \equiv \frac{1}{\sigma_\xi^2} B_c + \frac{1}{\sigma_q^2} \sum_{s=s_0}^{t-1} [q_{cs} - \Theta_S - \mu_{cs} k(\tau_{ct}) exp(-\alpha_{ct} p_{cs})]\omega_{cs} \tag{22}$$

where ϱ_j^{med}, $j = N, P$ are distributed according to the densities derived in Appendix 2. In equations (19) and (20) the sources of randomness are ε_{ct} and μ_{ct} respectively; on the other hand, equilibrium police enforcement (equation (21)) was derived as a deterministic function. While mean morality in the community is part of each individual's moral view, as an econometrician I can only estimate it. Thus, for estimation I will assume that z_{ct} is a normally distributed random variable with mean $\overline{z}_{ct}$ and variance σ_z^2: $z_{ct} \sim N(\overline{z}_{ct}, \sigma_z^2)$. Although at the individual level moral views are fixed over time (in the model this is actually also true at the dynasty level), average moral views in the city will vary as the demographic/religious distribution of the population changes. This is particularly relevant during the early decades of the Twentieth century, when both European immigration to the U.S. and internal migration to the West and from the South to the North were very dynamic. Because I will estimate mean moral views using observable heterogeneity (mainly the distribution of religious ascriptions), the stochastic component of this variable can be thought of as capturing measurement error, or any other sources of variation in average moral tastes for alcohol, which do not vary at the individual level (recall that individual-level moral shocks are unobservable, and incorporated in ϱ^i).

Given that the parameters of the model are identified only up to scale, I will normalize the variance of individual moral shocks ζ^i to 1. Interpretation of all other parameters will thus be relative to ζ^i.

I am interested in obtaining estimates of the parameters of this model, which will also allow me to directly compute estimates of the common component of belief sequences $\{\{\Omega_{ct}\overline{\theta}_{ct}^C\}_{t=1}^T\}_{c=1}^N$, and of the shape of the distribution of the median voter's unobserved preferred enforcement type ϱ_j^{med}. Parameters to be estimated are listed below:

Parameters	
Effect of Prohibition on crime	$\{\theta_c\}_{c=1}^N$
Alcohol market size	$\{a_c, b\}_{c=1}^N$
Law Enforcement	$\{\{\alpha_{ct}\}_{t=1}^T\}_{c=1}^N$
Collective Prior	$\{B_c\}_{c=1}^N$
State-specific crime shifter	$\{\Theta_S\}_{\forall S}$
Alcohol supply catch-up	κ, λ
Arrest probability	χ
Mean moral views	$\{\{\overline{z}_{ct}\}_{t=1}^T\}_{c=1}^N$
Variance of prior biases	σ_ξ^2
Variance of moral externality	σ_z^2
Correlation between moral views and prior biases	ρ

6.1 The Likelihood Function

I estimate the equilibrium political-economy model developed above through Conditional Maximum Likelihood (CMLE). Conditional on the decisive voter's ϱ_j^{med}, this economy is characterized by a system of equilibrium structural equations for crime, drunkenness arrests and police enforcement, plus an equation that pins down the learning dynamics of the common component of beliefs.

Individuals, who are assumed to know the model and its parameters, learn about θ_c by observing the realizations of the outcome vector $\boldsymbol{y}_{ct} = (p_{ct}, d_{ct}, q_{ct})$. The system in (19)-(21) has a particularly convenient "triangular" structure, which moreover, justifies the learning process implied by Bayesian learning and specified in equation (14). Once p_{ct} is realized, conditional on ϱ^{med} individuals face no uncertainty coming from equation (21) (recall that individuals observe z_{ct}). Then, after d_{ct} is realized, the realization of μ_{ct} can be exactly backed-up from equation (20). As a result, in equation (19) the only remaining uncertainty about crime comes from ε_{ct} and beliefs about θ_c, which is consistent with the conditional distribution of q_{ct} being normal, and hence, allowing the learning process to be as specified in section 5.

In Appendix 3 I derive the conditional likelihood function for the observed realization of the vector $\boldsymbol{y}_{ct} = (p_{ct}, d_{ct}, q_{ct})$. It is given by

$$\mathcal{L}_{ct}(\boldsymbol{y}_{ct}; \Theta_S, \theta_c, B_c, a_c, b, \alpha_{ct}, \chi, \overline{z}_{ct}, k, \lambda, \sigma_q^2, \sigma_z^2 | \varrho^{med}(P_{ct}), P_{ct}, \tau_{ct}) =$$

$$\frac{[g_\mu(\boldsymbol{y}_{ct})]^{a_c-1}[1-g_\mu(\boldsymbol{y}_{ct})]^{b-1}}{\int x^{a_c-1}(1-x)^{b-1}dx} \frac{1}{2\pi\sigma_q\sigma_z} exp(-\frac{1}{2\sigma_q^2} g_\varepsilon(\boldsymbol{y}_{ct})^2) exp(-\frac{1}{2\sigma_z^2}(g_z(\boldsymbol{y}_{ct}; \varrho^{med}(P_{ct})) - \overline{z}_{ct})^2) \frac{\partial g_\mu(\boldsymbol{y}_{ct})}{\partial d} \frac{\partial g_z(\boldsymbol{y}_{ct})}{\partial p} \quad (23)$$

where the expressions for $g_\varepsilon(\boldsymbol{y}_{ct})$, $g_\mu(\boldsymbol{y}_{ct})$, and $g_z(\boldsymbol{y}_{ct}; \varrho^{med}(P_{ct}))$ are given in the appendix. It is the product of a beta density coming from the distibution of the alcohol market size μ_t, two normal distributions coming from the shocks to the crime rate and the random variation in mean moral views, and the relevant jacobian of the transformation. Central to identification, discussed further below, the likelihood varies with P_{ct}. Prohibition introduces a structural change in the DGP, since a new nexus between law enforcement and criminality arises under Prohibition. A second key aspect of the model is that the DGP is dynamic; the vector of endogenous outcomes $\boldsymbol{y}_{ct}$ depends upon previous values of itself. In this model, the dynamic component comes, of course, from learning. The equilibrium choice of law enforcement at time t, p_{ct}, is a function of the current updated beliefs about θ_c, which depend on the whole sequence of previous realizations of the crime rate during Prohibition years $\{q_{cs}\}_{s=s_0}^{t-1}$. In the likelihood (equation (23)), the dynamic component enters through $g_z(\boldsymbol{y}_{ct}; \varrho^{med}(P_{ct}))$ exclusively.

While σ_q^2, σ_z^2, χ, κ, and λ are assumed constant across cities, I allow the parameters in the likelihood function above to vary with observable community characteristics as follows:

- $\Theta_S = \mathbf{x}_{\boldsymbol{S}}^{\boldsymbol{\Theta}\prime}\boldsymbol{\Sigma}$, where $\mathbf{x}_S^{\Theta}$ includes state-level dummies.
- $\theta_c = \mathbf{x}_{\boldsymbol{c}}^{\boldsymbol{\theta}\prime}\boldsymbol{\Lambda}$, where $\mathbf{x}_{\boldsymbol{c}}^{\boldsymbol{\theta}}$ includes border cities, South, state-capitals indicators, average demographics, and a constant.
- $B_c = \mathbf{x}_{\boldsymbol{c0}}^{\boldsymbol{B}\prime}\boldsymbol{\Xi}$, where $\mathbf{x}_{\boldsymbol{c0}}^{\boldsymbol{B}}$ is a vector containing the initial religious ascriptions distribution and a constant.
- $a_c = \overline{\mathbf{x}}_{\boldsymbol{c}}^{\boldsymbol{a}\prime}\boldsymbol{\Gamma}_a$ and b, where where $\overline{\mathbf{x}}_{\boldsymbol{c}}^{\boldsymbol{a}}$ includes average demographics, average religious ascriptions, average population, and a constant, and b is constant across cities[34].
- $\alpha_{ct} = \mathbf{x}_{\boldsymbol{ct}}^{\boldsymbol{e}\prime}\Psi$, where $\mathbf{x}_{\boldsymbol{ct}}^{\boldsymbol{e}}$ is a vector of legal enforcement variables (and a constant) such as the number of state-level dry laws in place (in years when the city is not under Prohibition these are the only source of restrictions on the alcohol market), a dummy equal to one when a city's state has a Prohibition enforcement law (during Prohibition), and other variables which might be correlated with federal law enforcement (a border city dummy, a Bureau of Prohibition period dummy, and dummies for the different Prohibition districts).
- $\overline{z}_{ct} = \mathbf{x}_{\boldsymbol{ct}}^{\boldsymbol{M}\prime}\boldsymbol{\Pi}$, where $\mathbf{x}_{ct}^{M}$ is a vector of containing the religious ascriptions distribution, and a constant.

Let $\boldsymbol{\beta} \equiv (\boldsymbol{\Sigma}, \boldsymbol{\Lambda}, \boldsymbol{\Xi}, \boldsymbol{\Gamma}_a, b, \chi, \kappa, \lambda, \boldsymbol{\Psi}, \boldsymbol{\Pi}, \sigma_q^2, \sigma_z^2)$, and $\mathbf{x}_{ct} \equiv (\mathbf{x}_{\boldsymbol{S}}^{\boldsymbol{\Theta}}, \mathbf{x}_c^{\theta}, \mathbf{x}_{\boldsymbol{c0}}^{\boldsymbol{B}}, \overline{\mathbf{x}}_{ct}^{a}, \mathbf{x}_{ct}^{M}, \mathbf{x}_{c0}^{B}, \mathbf{x}_{ct}^{e})$. The conditional likelihood can be more compactly written as $\mathcal{L}_{ct}(\boldsymbol{y}_{ct}; \boldsymbol{y}_{ct-1}, \mathbf{x}_{ct}, \boldsymbol{\beta}|\varrho_j^{med}, P_{ct}, \tau_{ct})$, which makes its dynamic nature explicit. Once the dynamic process is correctly specified (in this case the Bayesian learning assumption) and incorporated into the likelihood function, the density of the outcome vector $\boldsymbol{y}_{ct}$ only depends on $\boldsymbol{y}_{ct-1}$ through the learning channel, and hence the DGP is dynamically complete (See Wooldrige (2002, p. 412)). As a result, conditional on $\boldsymbol{y}_{ct-1}$, the $\boldsymbol{y}_{ct}$ are independently distributed. Thus, the conditional likelihood for a given observation $\boldsymbol{y}_c = (\boldsymbol{y}_{c1}, \boldsymbol{y}_{c2}, ...\boldsymbol{y}_{cT})'$ is given by $\mathcal{L}_c(\boldsymbol{y}_c, \boldsymbol{\beta}|\varrho^{med}, \boldsymbol{P}_c, \boldsymbol{\tau}_c) = \prod_t \mathcal{L}_{ct}(\boldsymbol{y}_{ct}; \boldsymbol{y}_{ct-1}, \mathbf{x}_{ct}, \boldsymbol{\beta}|\varrho^{med}(P_{ct}), P_{ct}, \tau_{ct})$, where ϱ^{med} is drawn

[34]While the first moment of the beta distribution is determined by the difference between a and b, its second moment is symmetrically decreasing in the magnitude of both a and b. Thus, allowing one of the parameters to depend on demographics and the religious distribution, while making the other one common across cities, allows this source of variation to identify the first and second moments. Allowing b to vary across cities could only increase the fit of the model. (This follows Coate and Conlin (2004)). Because I am assuming that a and b are constant across time for each city, I use the time-averaged values of the demographic and religious variables.

from $f_{\varrho_P^{med}}(\varrho^{med}; a_c, b, \sigma^2_\zeta, \sigma^2_\xi, \rho)$ during Prohibition years, and from $f_{\varrho_N^{med}}(\varrho^{med}; a_c, b, \sigma^2_\zeta)$ during years without Prohibition. Given that the ϱ_j^{med} are unobserved, it is necessary to integrate them out from the conditional likelihood, using their derived equilibrium densities. Estimates of $(\boldsymbol{\beta}, \sigma^2_\xi, \rho)$ are obtained from the following program:

$$max_{\boldsymbol{\beta},\sigma^2_\xi,\rho} \sum_c ln \left\{ \int_{-1}^{0} \left[\prod_t \mathcal{L}_{ct}(\boldsymbol{y}_{ct}; \boldsymbol{y}_{ct-1}, \mathbf{x}_{ct}, \boldsymbol{\beta} | \varrho^{med}(P_{ct})) \right] f_{\varrho^{med}(P_{ct})}(\varrho^{med}; a_c, b_c, \sigma^2_\xi, \rho, P_{ct}) d\varrho^{med} \right\} \quad (24)$$

As a final observation, dynamic models estimated by MLE usually face an "initial conditions" problem, arising from the fact that the observation for the first year in the sample depends upon an unobserved realization of the endogenous variable (See Wooldrige (2005)). In this model such a problem does not arise because for years under no Prohibition, the likelihood function does not depend on previous realizations of $\boldsymbol{y}_c$, and for the first period under Prohibition, the learning model implies that beliefs are exclusively based on the prior $\overline{\theta}_{c0}$, which is not a function of $\boldsymbol{y}_{ct-1}$ either. For all subsequent years under Prohibition, the relevant lagged information is available. Of course, this relies on having a sample covering for every observation, at least one year under no Prohibition.

Ideally, estimation of the model would cover the whole period; unfortunately, the drunkenness arrests data is only available for the years 1911-1929. Because this variable is necessary in the estimation to identify the alcohol market dynamics, I estimate the structural model for that period. Nevertheless, this imposes some discipline since it allows performing an out of sample exercise with the model's estimates to predict the observed data for the period 1930-1936. Thus, the sample used for the structural estimation consists of a fully balanced panel of 66 cities from 31 different U.S. states, for the nineteen year period 1911-1929. This makes a total of $1,254$ city-cross-year observations for which the homicide rate, the drunkenness arrest rate, the police expenditure share, and all of the demographic, religious and legal enforcement variables are available.

The only endogenous variable with a strong trend throughout the sample period, unaccounted for in the model, is the police expenditure share. Closer examination of the raw data reveals that this downwards trend is the result of a strongly increasing trend in total public spending across all cities in the United States during those years. Thus, for estimation I use the de-trended police expenditure share as the measure for p_{ct}[35]. As the crime outcome measure, I use the natural logarithm of the homicide rate, which standarizes the variance in homicide rates across cities, and is consistent with the shocks in equation (19) being normally distributed, and drawn from the same distribution across cities. Table A5-1 presents the list of cities included in the estimation and discusses the data further.

[35]While the average annual growth rate of total public spending in the sample was 5.6% (s.e. = 2.2%), the same number for police expenditure was only 3.7% (s.e.=2.5%). To obtain the detrended police share variable I ran a regression of the raw police expenditure share p^r_{ct} for each city in the sample, on a city-specific linear time-trend and city effects, and no constant: $p^r_{ct} = \alpha_c + \beta_c t + v_{ct}$. I then compute the detrended police share as $p_{ct} = \alpha_c + \widehat{v}_{ct}$. Of course, this is equivant to running a separate regression for each city.

6.2 Moments Identifying the Parameters in the Model

In this subsection I briefly discuss the relevant moments identifying the different parameters of the model. The structural elasticity of crime to the adoption of Prohibition, θ_c, is a function of city characteristics. Thus it is identified off the covariation in the homicide rate between cities with similar characteristics, and from the time-series variation in the homicide rate between periods under no Prohibition and periods under Prohibition. As previously noted, functional form is not key for the identification of θ_c, given that equation (19) can always be taken as a first order linear approximation to any monotonic relationship between the homicide rate and Prohibition enforcement.

Parameters a_c and b are identified off the residual variation in drunkenness arrests, once law enforcement and the catch-up of the alcohol supply have been accounted for. Since variation in law enforcement is correlated with variation in the availability of alcohol, the "wet" share cannot be identified from the drunkenness arrests data without additional information. This additional information comes from two sources: the variation in the homicide rate, by exploiting the fact that in a given city the drunkenness arrests and the homicide rate jointly covary with law enforcement, and the dynamics of the supply of alcohol under Prohibition, which the model assumes takes a particular functional form and is common across cities. It relies on two assumptions. First, that the baseline arrest probability, determined by χ, is constant over time, so that any changes in arrests between no-Prohibition and Prohibition years come solely from changes in law enforcement intensity, and not, for example, from changes in the "arrest technology". Second, that preferences over private alcohol consumption are independent of Prohibition status. Although a strong assumption in the context of Prohibition, a priori it is unclear in which direction tastes for alcohol might change when the community is under Prohibition. On the one hand, citizens might derive utility from abiding by the law, no matter what restrictions it imposes on their individual freedoms; on the other hand, they also could be subject to a "forbidden fruit" effect, where utility derived from a prohibited activity increases precisely because it is forbidden. Relatedly, since the baseline drunkenness arrest probability χ is assumed constant over time and across cities, χ is identified from the variation in arrest rates that is common across cities over time.

Regarding α_{ct}, the model assumes that the dynamics of the legal standard are exogenous from the point of view of the city. Although citizens were voting both for local law enforcement and for the state and federal legal standards, the assumption is that within a state or the Country as a whole, each city is too small to affect the equilibrium choice of legislation. This seems like a natural assumption, given that citizens in more rural areas were more strongly in favor of Prohibition. Indeed, many urban citizens of the United States saw the introduction of Prohibition as an intrusion from rural interests. Even in a state like New York, the pressure from Upstate voters set restrictions on the ability of New York City to dismantle Prohibition completely. At some level, this paper is about the effects of the imposition of a legal standard over communities where a large fraction of their members were in opposition to it. Thus, identification of α_{ct} comes from the common variation in drunkenness arrests and the homicide rate induced by changes in state-level legislation.

Identification of the city-specific collective prior, B_c, comes from early years under Prohibition, when the community choice of police enforcement closely follows prior beliefs. The larger the initial biases, the larger the gap between the observed police enforcement choice and what the optimal choice would be under perfect information. Because the model estimates θ_c, it implicitly provides a measure of how "off" law enforcement decisions were during early Prohibition years. In the model, the correlation of prior beliefs across cities depends on the distribution of religious ascriptions. Thus, the covariation between the gap from "optimal" law enforcement and the distribution of initial religious ascriptions identifies B_c.

On the other hand, the κ and λ parameters are identified from the common time-series residual variation in drunkenness arrests across cities, unaccounted for by changes in law enforcement or by changes in the wet share. The identification of these parameters relies strongly on the functional form I assume for the alcohol supply "catch-up" process, and the assumption that this catch-up is common for all cities in the sample. Nevertheless, the functional form in equation (8) is very flexible and can accommodate a wide variety of nonlinear trends.

Average moral views $\overline{z}$, which are function of the religious ascription distribution in the community are identified, from equation (21), from the variation in the police expenditure share which is uncorrelated with changes in beliefs, the alcohol market size, or dry legislation. Because the alcohol supply and beliefs change over time only during Prohibition years, the identification of $\overline{z}$ comes from the variation in law enforcement which is common for the city before and during Prohibition. On the other hand σ_z^2, the second moment of the distribution of z_{ct}, is identified directly from the sample variation in police enforcement that is common across cities.

Finally, σ_ξ^2 and ρ are identified in the model from the change in the shape of the estimated density of ϱ^{med} between no-Prohibition and Prohibition years, as figure 12 illustrates. As the variance of the distribution of biases decreases, the density of ϱ_P^{med} shifts to the left relative to the density of ϱ_N^{med}. This is because the weight on the prior is larger, and as a result, law enforcement choices give more weight, on average, to individuals' biases. The effect on the density of ϱ_P^{med} is similar as ρ increases in magnitude because a larger ρ (in absolute value) magnifies the differential law enforcement decision of dry cities relative to wet communities, increasing the variance of the distribution of ϱ_P^{med} relative to the distribution of ϱ_N^{med}. The reason is that if moral views ζ^i and belief biases ξ^i are correlated, this should have no effect on the preferences of the median voter when the city is not under Prohibition. During Prohibition, beliefs do shift the preferred police expenditure relative to no Prohibition periods, and the larger the correlation is (in absolute value), the larger the difference in the choice of optimal law enforcement between individuals with differing moral views. As ρ increases in magnitude, the density under Prohibition shifts mass to the left, making lower values of police expenditure more likely. Thus, ρ is of special interest in the estimation since it is identified off the channel stressed the most in Section 4 (the differential law enforcement choices between communities with varying moral views), highlighting the importance of the unobserved sources of heterogeneity in preferences over law enforcement for their dynamics during Prohibition.

6.3 Fit and Results

This section presents the estimation results from the CMLE. I start discussing the overall fit of the model's benchmark specification, and subsequently discuss the parameter estimates. To provide a general idea of the fit of the model across cities, panel A in figure 8 presents the city-level scatterplots of the average (over time) observed and predicted outcome variables, together with the

45 degree lines over which a perfect fit would obtain. The predictions are computed directly from equations (19)-(22), where I use the estimated expected value for the wet share μ_{ct} for each city, $a_c/(a_c+b)$, in the computation of the belief sequences, the predicted drunkenness arrest rate, and the predicted log homicide rate. For the predicted police shares, I use the mean value of the ϱ^{med}, which I calculate by integrating over the estimated equilibrium densities $f_{\varrho_P^{med}}(\varrho_P^{med}; a_c, b, \sigma_{\varrho Pt})$ and $f_{\varrho_N^{med}}(\varrho_N^{med}; a_c, b, \sigma_{\varrho N})$. The figures illustrate that the model does a fairly good job in fitting the variation across cities in the sample, especially for the drunkenness arrest rate and the homicide rate. The figure for the police share also shows a strong positive correlation ($= 0.44$), although the model tends to over-predict the small observed values and to under-predict large ones. The equation for the police share is no doubt the harder to fit, because preference heterogeneity, changes in the alcohol market size, and changes in beliefs are all interacting.

Regarding the time-series dimension, panel B in figure 8 presents the average (across cities) observed and predicted outcomes, for the sample years. For the three outcomes, the model is able to capture the joint evolution quite accurately, albeit with some differences in magnitudes. For example, it predicts a more pronounced fall in the police share than the one observed around the years 1920-1923, when the majority of cities were experiencing their first years under Prohibition. The apparent reason is that in the model, policing choices are quite sensitive to the size of the alcohol supply, and the impact effect of beliefs when cities enter into Prohibition is not large enough to counter the estimated fall in the alcohol supply. For the later years, the average predicted police share is around 0.1 percentage points larger than the observed. On the other hand, the predicted magnitude of the fall in the drunkenness arrest rate falls short from the one observed in the data between 1916 and 1920. The model is attributing a fraction of the fall in the drunkenness arrest rate to sampling variation from the distribution of the wet share μ_{ct}[36]. The model also predicts the fall to begin somewhat later, around 1918. Finally, the last figure depicts the predicted log homicide rate, showing that the model overpredicts the level of the homicide rate during the 1910s, and also predicts a smoother increase in this variable, compared to the rapid rise in homicides observed in the sample around 1920-1924. The reason for the overprediction of crime in early years is that I allow the alcohol market to have an effect on crime during the period without Prohibition. This suggests little or no room for an effect of the alcohol market on the homicide rate when Prohibition is not in place.

In addition, a way to asses the fit of the model is to look at the variability in the average moral views required to match the data. From equation (21), if the evolution of law enforcement, the alcohol supply, beliefs, and the change in the distribution of ϱ_j^{med} are able to match the police data closely, variation in average moral views z_{ct} over time should be small. In the model, the estimated $\sigma_z^2 = 0.06$ (s.e $= 0.03$), which relative to σ_ζ^2, the variance of individual moral tastes ζ^i (normalized to 1), is quite small. Overall, the estimates suggest that the mechanisms highlighted in the model

[36] The reason why the model predicts larger drunkenness arrests in the years in which these fall sharply is that by making the fall in the market supply larger, the fit of the police enforcement equation would be reduced because in the data, policing is not as sensitive to the dynamics of the alcohol supply. The large variability in drunkenness arrests could also be captured with a larger variance in the "wet share" distribution μ. Nevertheless, because the distribution of the data is positively skewed, an increase in the variance would require a larger estimate of a_c, which would imply a higher elasticity of the equilibrium police share to moral views and beliefs (see equation 21), reducing the ability of the model to fit the observed police outcomes.

capture a significant fraction of the joint variation in the data, despite the relatively small sample size.

6.3.1 Estimates

Estimates of the covariates from the model are presented in table 4, and table 5 presents the implied average estimated values of the main parameters of the model, based on the coefficient estimates. Standard Errors for the coefficients are computed through a bootstrap. Among the covariates for a_c, the demographic variables all have positive and significant coefficients as expected; cities with larger populations 15-44 years of age, larger foreign white populations, and larger black populations, have larger "wet shares". On the other hand, most of the coefficients on the religious ascriptions are unprecisely estimated, although the estimates for the religions traditionally considered as "wet" (Orthodox, Lutheran, and Catholic) are significant and positive. Finally, population size does not explain variation in the wet share. Together, the average estimate of a_c across cities is 3.16 and is 8.66 for b, implying that the average "wet share" is around 0.267. Since a_c varies little across cities (its standard deviation is 0.055), the model predicts very similar sizes of the "drinking population" across cities.

Looking at the covariates for average moral views $\overline{z}_c$, Baptist, Evangelical and Methodist shares do significantly increase average moral views. Surprisingly, the coefficient for the Catholic share is large in magnitude (0.81), but imprecisely estimated. Looking at the covariates for α_{ct}, the alcohol-related laws variable is insignificant (point estimate = 0.11, s.e.= 0.25), suggesting that changes in dry legislation had little effect in making policing more effective for Prohibition enforcement. On the other hand, the coefficient on the Enforcement Law dummy is negative and quite significant (point estimate = -0.89, s.e.= 0.21), suggesting that the repeal of state-level Prohibition enforcement laws made policing more effective for crime enforcement. This might be driven by the unwillingness of local authorities to enforce Prohibition laws which they oppose. Regarding the Prohibition Unit indicators, out of which the New York Unit is omitted, all other Units except for the San Francisco and Los Angeles ones have negative estimated coefficients. This is consistent with historical observation that federal law enforcement was especially focused around the mid-Atlantic "wet" states, and with a relatively dry coastal California, which likely made a given amount of policing more effective in reducing the alcohol market.

Of central interest are the model's estimates of θ_c, the structural "elasticity" of Prohibition enforcement to crime. The average θ_c is 1.37, and figure 9 presents the distribution of estimated θ_c's across cities. These range from around 0.8 to 1.6. At the means of the police share p_{ct} and the estimated parameters, it implies that the average city saw an increase in the homicide rate of around 23% during Prohibition[37]. Among the estimates for its covariates, the estimate for the border indicator (Canadian, Mexican or coastal city) is negative and significant (point estimate = -0.302, s.e.= 0.101). Given the accounts of huge amount of smuggling during the Prohibition years, this is at first puzzling, but actually consistent with my discussion above, about borders having the effect

[37]The average (normalized) police share is 0.36. Assuming an 80% size of the alcohol supply (around the 9th year under Prohibition using the estimated κ and λ), and using the mean estimate of $\alpha_{ct} = 3.43$, $a_c/(a_c + b) = 0.267$ and $\theta_c = 1.37$, it follows that $0.23 = exp(1.37 \times 0.8 \times 0.267 \times [1 - exp(-3.43 \times 0.36)]) - 1$.

of increasing the availability of alcohol for a given level of law enforcement, and thus, reducing the incentives for Prohibition-related crime to arise.

Estimates for the covariates of the Prior B_c are also presented in table 4. Baptist and Methodist shares have the largest (in magnitude) estimated significant coefficients, implying that cities with larger fractions of members of these religions initially did have more optimistic Priors about the effects Prohibition would bring. The Catholic share also has a coefficient of large magnitude, but once again its standard error is quite large. On the other hand, the estimates for the second moments of the joint distribution of individual biases and moral views (see equation (6)) also are of interest. The variance of biases σ_ξ^2 is estimated to be 0.34, implying that the variation in individual moral views (recall its variance σ_ζ^2 was normalized to 1) was significantly larger than variation in biases. In the model, the magnitude of σ_ξ^2 is a measure of how much weight people put on their prior beliefs, so that smaller values of σ_ξ^2 directly imply slower learning. Finally, ρ, the estimated correlation between prior biases (ξ^i) and moral views (ζ^i) is -0.49 (s.e.$=0.9$), suggesting that cities with constituencies more favorable to Prohibition did have much more optimistic beliefs about its effects.

An alternative way to see the correlation between moral views and beliefs from the model's estimates is with a scatterplot of the estimated values of priors B_c and average moral views $\overline{z}_c$, for the cities in the sample. Figure 10 presents such a scatterplot, together with a simple regression line. Its slope is -0.25 with a t-statistic of -6.36. Thus, even in this sample of relatively large cities, average prior beliefs and moral views were negatively correlated. In particular, the model predicts negative values of prior beliefs for all cities in the sample. This is because the cities observed a relative increase in policing in the early years under Prohibition (see figure 8), which in the model is driven by optimistic priors. Because early on during Prohibition learning is slow, the sharp fall in the alcohol supply more than offsets the average increase in beliefs, explaining the subsequent fall in policing observed in the data.

The parameter estimates from table 5 also allow a quantitative characterization of the structural relationships specified in the model. In particular, the estimates for κ and λ from equation (8) (0.26 and 0.25) imply that at its lowest point, the supply of alcohol was on average 68% its pre-Prohibition level, and that this minimum was attained around 3.75 years after the introduction of Prohibition[38]. Together with this estimated function for the alcohol supply catch up, figure 11 presents the estimated drunkenness arrest conditional probability, and the estimated percent increase in the homicide rate due to Prohibition, both as a function of police expenditure[39]. The three graphs in the figure present an illustrative picture of the costs and benefits of Prohibition. Prohibition was able to shrink the alcohol supply by about 35%, but only for a relatively short period of time. While increasing policing would increase arrests for drunkenness, the slope is not very steep. Considering that the average standard deviation of (normalized) police shares in the sample is 0.044, a whole standard deviation increase in the police share would at most increase the arrest probability by 3%. In sharp contrast, the same increase in policing during Prohibition would imply that the homicide rate would move from being 23 to 24.6% higher under Prohibition[40].

[38]The minumim of equation 8 is attained at $\tau = 1/\kappa$.

[39]Thus, using the average parameter estimates from table 5, the estimated arrest probability is computed as $Pr(Arrest|p) = \frac{exp(3.43p)}{8.94+exp(3.43p)}$, and the estimated proportional increase in the homicide rate under Prohibition is computed as $\Delta Q(p) = exp(1.37 \times 0.8 \times 0.267 \times [1 - exp(-3.43p)])$, for $k(\tau) = 0.8$.

[40]Average (normalized) police share is around $p = 0.36$. Thus, the increase in the arrest probability induced

Finally, the estimated shapes of the distributions of the unobserved ϱ_j^{med} (the median voter's "type" under no Prohibition and under Prohibition) can be directly derived from the parameter estimates of the structural model, by plugging the estimates of a_c, b, σ_ξ^2, σ_ζ^2, and ρ in equation (32) from Appendix 2. Figure 12 plots both densities, for the mean values of the parameter estimates, and for the first year under Prohibition (when $\Omega_t = \sigma_\xi^2$) . The difference in skewness between the distribution under Prohibition and under no Prohibition is what identifies ρ in the model. This is because the larger (in magnitude) the correlation between moral views and belief biases, the larger the average difference in policing choices that a median voter would make, when passing from no Prohibition to Prohibition. Also, as t increases, $\Omega_t \to 0$, so that the density under Prohibition converges to the density under no Prohibition.

6.3.2 Learning

In this subsection I discuss the estimation results related to learning. Recall the model estimates a relatively low variance of individual belief biases σ_ξ^2. This is the main exogenous parameter affecting the speed of learning in the model, and is common across cities. Thus, differences across cities in the estimated speeds of learning are due directly to the variation in enforcement choices over time, which under normal updating, affect the informativeness of the signal. The reason for the relatively low estimate of σ_ξ^2 is that the model is estimated over the years 1911-1929, thus, excluding precisely the later years under Prohibition (1930-1936), in which the largest adjustments in law enforcement occured. Nevertheless, there is substantial learning over the nineteen year period. Figure 13 graphs the evolution of the estimated empirical distribution of the common component of beliefs $\{\{\Omega_{ct}\overline{\theta}_{ct}^C\}_{c=1}^N\}_{t=1911}^{1929}$, derived directly from applying equation (22) iteratively using the estimated coefficients and the observed sequences of outcome variables. The outermost curves represent the 10th and 90th percentiles, the curves in between represent the 25th and 75th percentiles, and the middle curve represents the median of the estimated distribution. Of course, beliefs remain at the prior until cities fall under Prohibition status. In several cities, for some of the early Prohibition years, beliefs about θ_c actually fall slightly. After around 1923 though, the belief sequences are monotonically increasing for all cities, but there is substantial variation in the speed of belief updating. The figure also shows that despite the generalized increasing pessimism over the effects of Prohibition, the dispersion of beliefs actually increases over time. Mean common beliefs increase from the average prior $B_c = -1.31$ to a mean posterior of -0.57 in 1929, whereas the posterior median is only around -0.75. While the standard deviation of priors is 0.26, it is 0.68 for the 1929 posteriors. At some level, this is a natural implication of the model, given that each city is learning from its own experience exclusively, and that different cities had different structural values of θ_c.

A key question is whether the differential evolution of beliefs across cities, is correlated with differences in their moral profiles. The reduced-form estimates already suggested that this is the case. Recall from Section 4 that during the first years under Prohibition, wetter cities had differentially lower levels of police enforcement. I argued there that this could be driven by the willingness of

from increasing policing to $0.4 = 0.36 + 0.04$ would be $\frac{exp(3.43\times0.4)}{8.94+exp(3.43\times0.4)} - \frac{exp(3.43\times0.36)}{8.94+exp(3.43\times0.36)} = 0.028$. On the other hand, the shift in the homicide rate goes from $exp(1.37 \times 0.8 \times 0.267 \times [1 - exp(-3.43 \times 0.36)]) - 1 = 0.232$ to $exp(1.37\times0.8\times0.267\times[1-exp(-3.43\times0.4)])-1 = 0.246$ under Prohibition, by increasing policing by one standard deviation around its average.

more optimistic “dry” cities to invest in law enforcement. In the context of a learning model, dry cities should learn faster early on, given that their signals are more precise. The estimates here are consistent with that view; running a regression of the estimated 1929 posteriors on the estimated average moral views, and controlling for the estimated priors, the coefficient estimate on moral views is positive and has a t-statistic of 2.26[41]. Thus, although the standard deviation of beliefs across cities increased over time, the incentives for differentially higher initial law enforcement in drier cities limited the extent of divergence in beliefs. This is also consistent with the fact that among the subset of relatively “wetter” communities, dry ones saw larger shifts of public opinion against Prohibition (see figure 7). Overall, the structural estimates of the evolution of beliefs are consistent with the correlations from the reduced-form analysis.

At the heart of the model is the endogenous evolution of outcomes due to rational learning. Thus, I end this subsection by estimating the model closing the learning channel, to assess the relative performance of a model where no learning occurs compared to the benchmark specification (this follows Buera et al. (2010)). Formally, this is equivalent to imposing the restriction $\sigma_\xi^2 = 0$, so that individuals never update their priors. A simple Likelihood Ratio test can be performed comparing the restricted No-Learning model with the benchmark model. The log-likelihood for the model without learning is $5,978.99$, while the log-likelihood for the benchmark model is $6,560.77$. Under the null hypothesis that the restricted and unrestricted models are indistinguishable,

$$LR = 2[logL(Benchmark) - logL(NoLearning)] \sim \chi^2_{701} \tag{25}$$

Assuming $\sigma_\xi^2 = 0$ implies a restriction in the police equation for each city, in every year under Prohibition except the first. There are 767 such observations, so the appropriate number of degrees of freedom for the test's χ^2 distribution is 701. While $LR = 1,163.55$, the 99% critical value is 791.03. Thus, the null can be rejected at any significance level.

6.4 Counterfactuals

To conclude, I exploit the model's estimates to perform a series of counterfactual exercises. These should allow a further assessment of the model's fit, and also provide general-equilibrium answers to questions of interest, which would be impossible to make in a partial equilibrium or reduced-form framework. First, I perform an “out of sample” prediction of the outcome variables for the years 1930-1936, using the parameter estimates. I then ask the following questions to the model: what would the evolution of outcomes have been under Prohibition, if average prior beliefs had been unbiased? How would Prohibition outcomes have evolved if society had been more radicalized? More polarized? Finally, I assess the implications of alternative political environments.

[41] For the 66 cities in the sample, I run the regression $\Omega_{c,1929}\overline{\theta}^C_{c,1929} = \beta_0 + \beta_1 \overline{z}_c + \beta_2 \Omega_{c,1911}\overline{\theta}^C_{c,1911} + \varepsilon_c$. The estimated β_1 is 2.17 with a standard error of 0.96. I include the prior as a regressor to control for the fact that morally drier cities had more negative priors.

6.4.1 Out of Sample Prediction

Because of the unavailability of drunkenness arrests data for years after 1929, I am unable to estimate the model for the later Prohibition years. Thus, I make an out of sample prediction for the police and homicide outcomes during the 1930-1936 years, by using the MLE estimates on equations (19)-(22). This exercise is particularly meaningful because I do observe the police and homicide rate outcomes in that period, so I directly can assess the extent to which the model is able to capture the subsequent evolution of Prohibition during its final phase, and the first few years after its repeal. For this purpose, I take the estimated 1929 posterior beliefs for each city, and use them as the 1930 priors. I then compute iteratively the predicted equilibrium values of p_{ct} from equation (21), and with this predicted police enforcement value, I then predict q_{ct} from equation (19). To compute year t's posterior from equation (22), I add a random shock drawn from a mean-zero normal distribution with variance equal to 0.277 (the MLE estimate for the variance of ε, σ_q^2) to the predicted value of q_{ct} and iteratively use this posterior to calculate year $t+1$'s police choice and homicide rate. Constitutional Prohibition was repealed in the end of 1933, so belief updating actually stops after this year. Figure 14 presents graphs analogous to those in figure 8, comparing the "out of sample" average predicted values from the structural model, both in the time and city dimensions. Panel A shows the scatterplots of the 1930-1936 averages for each city. The horizontal axis has the observed values, while the vertical axis has the predicted values. The predictions for the homicide rate are again fairly close to the observed. For the policing data, the slope is significantly positive, but as in the predictions for the 1911-1929 period, the model is not able to capture all of the variablility across cities. Looking at panel B, on the other hand, it captures the trend of both variables over time remarkably well, in particular the fall in both policing and the homicide rate during the last years of Prohibition, and the leveling off of both variables after repeal.

6.4.2 Changes in Prior Beliefs

The adoption of Prohibition in the U.S. would not have been possible based exclusively on moral motivations, since radically dry sectors did not constitute a large enough majority of the population. Its adoption required a large fraction of morally-indifferent voters to have optimistic beliefs about the effects of the policy. Thus, a natural question arises: what was the cost of these biased prior beliefs? The model can provide an answer to this question, by making the counterfactual exercise of assuming that priors were unbiased. Specifically, I assume that prior common beliefs in 1911 were unbiased, this is, that $B_c = \theta_c$. Using the estimated coefficients, I can then compute the predicted evolution of outcomes over time, and compare them to the model's predicted outcomes under the estimated biased priors.

The simulation results reveal several patterns. As expected, beliefs endogenously remain fairly unchanged over time, since the realized homicide outcomes are always close to the expected ones given the law enforcement choices. Police enforcement decisions, on the other hand, behave quite differently. In particular, cities would have avoided the early-Prohibition increases in policing, since in the absence of optimism about Prohibition's effects, there are no incentives to increase law enforcement. Subsequently, policing decisions would have fallen sharply relative to the benchmark case, following the early contraction of the alcohol supply, and would have bounced back at a relatively faster pace. In contrast, when beliefs are biased, learning makes this effect nuanced as the median voter finds it less attractive over time to maintain high levels of police enforcement. The model predicts that the median city would have reduced law enforcement to almost half the predicted law enforcement levels under biased beliefs. Thus, cities would have been much more radical in offsetting Prohibition with their local law enforcement choices. Variation across cities in law enforcement would have increased, on the other hand, because the variance in the distribution of Prohibition-related crime potential θ_c is larger than the variation in estimated priors. In addition, the model also suggests that the differences in the homicide rate relative to the biased-beliefs case would have been insignificant. This is because the inability to reduce Prohibition enforcement without concomitantly reducing overall crime enforcement implies that the relatively large fall in policing would allow for an increase in non-Prohibition related crime. Somewhat counterintuitively, this suggests that conditional on Prohibition been imposed, more accurate initial beliefs about its effects could have allowed the policy to remain in place longer, because large cities would have faced relatively similar crime outcomes, but lower police enforcement expenditures. Since beliefs would not have changed significantly, public opinion would have been limited.

6.4.3 Radicalization and Polarization

The model also can address questions related to the distribution of preferences in society. I perform two simple exercises. I start by asking what the evolution of outcomes under Prohibition would have been, relative to the estimated benchmark model, under more radical moral views against alcohol consumption. This implies a higher degree of alignment between the prohibitive legal standard and preferences over its enforcement. Consequently, I increase each city's estimated average moral view $\overline{z}_{ct}$ by one or two standard deviations (the estimated $\sigma_z^2 = 0.06$), and compute the predicted sequences of outcomes under these changes. The model predicts that these radicalized communities would choose a constantly higher level of police enforcement (around 20% more for the one standard deviation increase, and around 36% more for the two standard deviation increase), but variation in police choices across cities would also be larger. Common beliefs would consequently be updated faster relative to the benchmark model's predictions. Nevertheless, in this case it is unclear whether public opinion would turn against Prohibition as fast as it in fact did, given that across cities the decisive voter is more willing to restrict the alcohol market for a given belief profile.

Another exercise of interest is to increase the degree of polarization in society. By polarization here I mean increasing the average willingness to enforce prohibition, by raising $\overline{z}_{ct}$, and at the same time increasing the demand for alcohol, by raising the mean of the distribution of μ. Thus, just as in the counterfactual exercise above, I allow $\overline{z}_{ct}$, but also $E[\mu]$, to increase by one or two standard deviations. The estimated standard deviation of the "wet share" μ is 0.123, while its mean is 0.267, so that a one standard deviation increase in the mean implies $E[\mu] = 0.38$. Holding a fixed, such a shift in the distribution of the wet share can be achieved by reducing the value of b to 5. For a two standard deviation increase in the mean of μ, which implies $E[\mu] = 0.5$, a value of $b = 3.1$ achieves the same objective. The model predicts that the speed of learning during Prohibition increases very fast on the degree of polarization in society. The benchmark model's estimated 1929 posterior common beliefs for the median city would have been reached by 1923 if both average moral views and the average wet share were one standard deviation larger, and by 1921 if they were two standard deviations larger. This outcome is the result of increased police enforcement levels as the degree of polarization increases. Given the model's parameter estimates, this occurs for two reasons. First, more radical moral views increase the ideal choice of Prohibition enforcement across the population. Moreover, because prior beliefs were initially relatively optimistic, a larger wet share also gives incentives for the median voter to prefer more law enforcement, since the perceived moral externality is larger for a given moral view, while the expected cost of increased crime is low.

On the other hand, policing choices would have been much more stable over time because the increased salience of the moral externality reduces the extent to which the police expenditure responds to changes in the alcohol supply. Nevertheless, as an added equilibrium effect, the distribution of police enforcement choices across cities spreads out considerably. The apparent reason is a political economy effect; because a larger wet share shifts the median voter towards "wetness", there is a force driving the equilibrium choice of law enforcement downwards. Finally, the model predicts that these polarized communities would observe significantly higher levels of crime during Prohibition. For instance, the median city would have on average 2.9 more homicides per hundred thousand on the average Prohibition year in the two standard deviations higher polarization society, or 1.37 more in the one standard deviation higher polarization case. Thus, although communities with more extreme preference distributions do learn much faster about the structural relationship between Prohibition and crime, they also face a constituency much more willing to endure the increased levels of crime.

6.4.4 Alternative Political Environments

In the setting of this model, it is natural to ask what would the equilibrium effects of changes in the political environment be. This is important because, as I have shown, the equilibrium collective law enforcement decisions play a central role in the success or failure of a given legal standard. In particular, I ask about the effect of interest groups in politics by assuming that some constituencies have more political power than others, shifting the decisive voter away from the median. To make the intiutions clear I look at the polar cases in which the decisive voter in the community is either the median voter among the wets (the decisive voter's type is $\varrho_j = -1$), or the median voter among the drys (the decisive voter's type is $\varrho_j = 0$). Under each conterfactual scenario I compute the predicted outcome sequences, using the benchmark parameter estimates.

Results are closely related to the ones above. When drys have all the political power, law enforcement chioces are consistently larger in magnitude relative to the benchmark case. Because alcohol demand remains unchanged, these enforcement choices increase the informativeness of the crime signals, making beliefs evolve faster. Belief sequences across the distribution of cities under this counterfactual scenario are on avergage two years ahead relative to the benchmark case. Consistently, the predictions of the counterfactual simulation where wets have all the political power deliver weaker law enforcement relative to the benchmark, which consequently translates into slower learning. The benchmark estimated average beliefs in 1925 would only be reached in 1929 under this counterfactual setting.

These results are driven by the increased divergence between the decisive voter over law enforcement and average voters' preferences. They make the point that the effects of increased conflict also arise when the identity of those deciding over law enforcement is further away from overall constituency preferences. Increased conflict, in this setting due to a skewed collective decisionmaking process, is a force driving changes in public opinion. When drys have all the politcal power at the local level, their choices of law enforcement are too large relative to what the community's median voter would prefer, and relative to the community's alcohol market size. As a result, crime outcomes are more informative and communities learn faster. In the polar opposite case, when all political power is allocated to the wets, learning is too slow relative to the benchmark because the very weak law enforcement choices make crime realizations uninformative.

7 Conclusions

Many central political cleavages in contemporary societies revolve around ideological or moral issues, over which people frequently have strong and polarized views. I have highlighted learning about policies, and the endogenous dynamic feedback between enforcement choices and policy support, as a driving force for changes in public opinion over moral issues, and more broadly for social change, by looking at the U.S. Prohibition experience during the early decades of the Twentieth century. The circumstances around Prohibition were very specific to that policy; in particular, the potential effects that closing the alcohol market could have over crime are very specific to prohibitions. Nevertheless, looking at the side-effects (or absence thereof) of policies, and at learning about them, can allow a better understanding of the evolution of policy reform over social cleavages. The extent to which people are informed is important, and of course, the political economy of the extent of such information acquisition becomes key; this should be an area of future research.

I developed a model of endogenous learning and law enforcement in a political economy framework, which has some success in replicating the patterns observed in the data. The paper suggests that an important element to understand the effects and success of policies is the degree of alignment of the legal standard and the law enforcement choices associated with it. This was particularly relevant during Prohibition because most of the law enforcement was decided at the local level, while the prohibitionist legal standard was chosen either at the state or nationwide levels. The estimates suggests that prior beliefs about Prohibition's effect on crime were very optimistic and highly correlated with moral views, that local policy responded closely to communities' preferences, and that community preferences also were responsive to changes in beliefs. In the model, the estimated speed of learning is relatively slow. This might be due to the assumption of exclusively localized learning, whereas it is likely that individuals' opinions were also affected by outcomes across the country. Learning from neighboring communities is likely to be important in societies where the media plays a large role in shaping public opinion. This constitutes an avenue for improvement of the structural model, and for understanding other instances of social change. This paper did not exploit the judicial dimension of law enforcement either, although prohibition enforcement at the local level was also implemented through judicial prosecution. Further research should look at the evolution of judicial decision-making regarding Prohibition as an alternative law enforcement mechanism, which was likely subject to different political economy incentives.

References

Aghion, Phillipe, Yann Algan, Pierre Cahuc, and Andrei Schleifer, "Regulation and Distrust," 2008. Mimeo, Harvard University.

Alesina, Alberto and Edward Glaeser, *Fighting Poverty in the US and Europe: A World of Difference*, Oxford, UK: Oxford University Press, Forthcoming.

_ **and Nicola Fuchs-Schundeln**, "Good-Bye Lenin (or Not?) The Effect of Communism on People's Preferences," *American Economic Review*, 2007, *97* (4), 1507–1528.

Arrow, Kenneth, *Social Choice and Individual Values*, New Haven: Cowles Foundation, 1963.

Asbury, Herbert, *The Great Illusion. An Informal History of Prohibition*, Garden City, NY: Doubleday and Co., 1950.

Baland, Jean Marie and James Robinson, "Land and Power: Theory and Evidence from Chile," *American Economic Review*, December 2008, *98* (5), 1737–1765.

Baron, David, "Electoral Competition with Informed and Uninformed Voters," *The American Political Science Review*, 1994, *88* (1), 33–47.

Becker, Gary, "Crime and Punishment," *The Journal of Political Economy*, March 1968, *76* (2), 169–217.

Beman, Lamar T., *Prohibition. Modification of the Volstead Act*, The H.W. Wilson Company, 1924.

Blocker, Jack, *Retreat from Reform. The Prohibition Movement in the United States, 1890-1913*, Westport, CT: Greenwood Press, 1976.

_ , *American Temperance Movements. Cycles of Reform*, Boston, MA: Twayne Publishers, 1989.

_ , "Did Prohibition Really Work? Alcohol Prohibition as a Public Health Innovation," *Public Health Then and Now*, 2006, *96* (2), 233–243.

Brown, Everett, "The Ratification of the Twenty-First Ammendment," *The American Political Science Review*, December 1935, *29* (6), 1005–1017.

Buera, Francisco, Alexander Mongue-Naranjo, and Giorgio Primiceri, "Learning the Weath of Nations," 2010. Unpublished.

Capper, Arthur, "Politics in the Enforcement of the Liqour Laws," *Annals of the American Academy of Political and Social Science*, 1923, *109*, 155–164.

Carpenter, Christopher and Carlos Dobkin, "Alcohol Regulation and Crime," 2010. NBER Working Paper No. 15828.

Chaudrhi, Vivek and John Geanakopolos, "A Note on the Economic Rationalization of Gun Control," *Economic Letters*, 1998, *58*, 51–53.

Cherrington, Ernest H., *The Evolution of Prohibition in the United States of America*, Westerville, OH: The American Issue Press, 1920.

Clark, Norman, *The Dry Years. Prohibition and Social Change in Washington*, Seattle: University of Washington Press, 1965.

Coate, Stephen and Michael Conlin, "A Group Rule-Utilitarian Approach to Voter Turnout: Theory and Evidence," *American Economic Review*, 2004, *94* (5), 1476–1504.

_ , _ , and Andrea Moro, "The Performance of Pivotal-voter models in small-scale elections: Evidence from Texas Liquor Referenda," *Journal of Public Economics*, 2008, *92*, 582–596.

Cochran, John and Mitchell Chamlin, "Can Information Change Public Opinion? Another Test of the Marshall Hypothesis," *Journal of Crime Justice*, 2005, *33* (6), 573–584.

Colvin, Leigh, *Prohibition in the United States*, New York: George H. Doran Company, 1926.

DalBo, Ernesto, Pedro DalBo, and Rafael DiTella, "Plata o Plomo?: Bribe and Punishment in a Theory of Political Influence," *American Political Science Review*, 2006, *100* (1), 41–53.

Degan, Arianna and Antonio Merlo, "A Structural Model of Turnout and Voting in Multiple Elections," March 2009.

DeGroot, Morris, *Optimal Statistical Decisions*, McGraw-Hill Book Company, 1970.

Dills, Angela, Jeffrey Miron, and Garrett Summers, "What do Economists Know About Crime?," January 2008. NBER Working Paper No. 13759.

_ , Mireille Jacobs, and Jeffrey Miron, "The Effect of Alcohol Prohibition on Alcohol Consumption: Evidence from Drunkenness Arrests," *Economics Letters*, 2005, *86* (2), 279–284.

DiTell, Rafael and Robert MacCulloch, "Why Doesn't Capitalism Flow to Poor Countries?," June 2007. NBER Working Paper No. 13164.

Donohue, John and Steve Levitt, "The Impact of Legalized Abortion on Crime," *The Quarterly Journal of Economics*, 2001, *116* (2), 379–420.

Duflo, Esther, Rema Hanna, and Stephen Ryan, "Incentives Work: Getting Teachers to Come to School," May 2010. Unpublished.

Eckberg, Douglas Lee, "Estimates of Early Twentieth-Century U.S. Homicide Rates: an Econometric Forecasting Approach," *Demography*, 1995, *32* (1), 1–16.

Edge, Walter, "The Non-Effectiveness of the Volstead Act," *Annals of the American Academy of Political and Social Science*, 1923, *109*, 67–84.

FBI, *Uniform Crime Statistics*, Vol. 1.1, 1.5, 2.12, 5.4, 6.4, 7.4, 8.2 and 8.3, United States Government Printing Office, 1930-1939.

Fisher, Irving, *Prohibition Still at its Worse*, New York: Alcohol Information Comittee, 1928.

Foster, Gaines, *Moral Reconstruction: Christian Lobbyists and the Federal Legislation of Morality, 1865-1920*, The University of Carolina Press, 2002.

Frankfurter, Felix, "A National Policy for Enforcement of Prohibition," *Annals of the American Academy of Political and Social Science*, 1923, *109*, 193–195.

Franklin, Fabian, "What's Wrong with the Eighteenth Amendment?," *Annals of the American Academy of Political and Social Science*, 1923, *109*, 48–51.

Franklin, Jimmie Lewis, *Born Sober: Prohibition in Oklahoma, 1907-1959*, Norman, OK: University of Oklahoma Press, 1971.

Goldstein, Paul, *The Drugs/Violence Nexus: A Tripartite Conceptual Framework*, 1985, *39*, 143–174.

Haider-Markel, Donald and Kenneth Meier, "The Politics of Gay and Lesbian Rights: Expanding the Scope of the Conflict," *The Journal of Politics*, May 1996, *58* (2), 332–349.

Harrison, Leonard and Elizabeth Laine, *After Repeal. A Study of Liquor Control Administration*, New York: Harper and Brothers, 1936.

Hayek, Frederick, *The Constitution of Liberty*, Chicago, IL: The Chicago University Press, 1960.

Heckman, James, "Building Bridges Between Structural and Program Evaluation Approaches to Evaluating Policy," June 2010. NBER Working Paper No. 16110.

Hobart, George, "The Volstead Act," *Annals of the American Academy of Political and Social Science*, 1923, *109*, 85–101.

Hohner, Robert, "Prohibition comes to Virginia. The Referendum on 1914," *The Virginia Magazine of History and Biography*, October 1967, *75* (4), 473–488.

Holmes, Thomas, "The Effect of State Policies on the Location of Industry: Evidence from State Boundaries," September 1996. Federal Reserve Bank of Minneapolis.

Isaac, Paul, *Prohibition and Politics: Turbulent Decades in Tennessee 1885-1920*, Knoxville, TN: The University of Tennessee Press, 1965.

Jackson, Joy, "Prohibition in New Orleans: The Unlikeliest Crusade," *Louisiana History: The Journal of the Louisiana Historical Association*, 1978, *19* (3), 261–284.

Kamada, Yuichiro and Fuhito Kojima, "Voter Preferences, Polarization and Electoral Policies," January 2010. Unpublished.

Koren, John, "Inherent Frailties of Prohibition," *Annals of the American Academy of Political and Social Science*, 1923, *109*, 52–61.

Kyvig, David, *Repealing National Prohibition*, Chicago: The University of Chicago Press, 1979.

Lagunoff, Roger, "A Theory of Constitutional Standards and Civil Liberties," *Review of Economic Studes*, 2001, *68* (1), 109–132.

Landier, Augustin, David Thesmar, and Mathias Thoening, "Investigating Capitalism Aversion," *Economic Policy*, 2008, *23* (55), 465–497.

Lax, Jeffrey and Justin Phillips, "Gay Rights in the States: Public Opinion and Policy Responsiveness," *American Political Science Review*, 2010, *Forthcoming.*

League, Anti-Saloon, *Anti-Saloon League Yearbook*, Anti-Saloon League of America, 1932.

Levitt, Steven, "Understanding Why Crime Fell in the 1990s: Four Factors that Explain the Decline and Six that Do Not," *The Journal of Economic Perspectives*, 2004, *18* (1), 163–190.

Lewis, Michael, "Accounting for Differences in Local and State Alcohol Laws, North Carolina in 1908," NA.

Martis, Kenneth, *The Historical Atlas of U.S. Congressional Districts, 1789-1983*, Free Press, 1982.

Miron, Jeffrey, "The Effect of Alcohol Prohibition on Alcohol Consumption," 1999. NBER Working Paper No. 7130.

_ , "Violence and the U.S. Prohibitions of Drugs and Alcohol," *Americal Law and Economics Review*, 1999, *1* (1), 78–114.

_ **and Jeffrey Zweibel**, "Alcohol Consumption during Prohibition," *American Economic Review*, 1991, *81* (2), 242–247.

Mukand, Sharun and Dani Rodrik, "In Search of the Holy Grail: Policy Convergence, Experimentation and Economic Performance," *American Economic Review*, 2005, *95* (1), 374–383.

Okrent, Daniel, *Great Fortune: The Epic of Rockefeller Center*, New York: Viking Press, 2003.

_ , *Last Call: The Rise and Fall of Prohibition*, New York: Scribner, 2010.

Polinsky, Michael and Steven Shavell, "The Theory of Public Enforcement of Law," in Michael Polinsky and Steven Shavell, eds., *The Handbook of Law and Economics*, Vol. 1, Elsevier, 2007, chapter 6.

Polinsky, Mitchell and Steven Shavell, "Corruption and Optimal Law Enforcement," *Journal of Public Economics*, 2001, *81* (1), 1–24.

Razin, Ronny, "Signaling and Election Motivations in a Voting Model with Common Values and Responsive Candidates," *Econometrica*, July 2003, *71* (4), 1083–1119.

Sargent, Thomas, Noah Williams, and Tao Zha, "Shocks and Government Beliefs: The Rise and Fall of American Inflation," *American Economic Review*, 2006, *96* (4), 1193–1224.

Schaller, Thomas, "Democracy at Rest: Strategic Ratification of the Twenty-First Amendment," *Publius: The Journal of Federalism*, 1998, *28* (2), 81–97.

Schmeckebier, Laurence, *The Bureau of Prohibition. Its History, Activities and Organization*, Washington D.C: The Brookings Institution, 1929.

Schmekebier, Laurence, *The Bureau of Internal Revenue. Its History, Activities and Organization.*, Baltimore: The Johns Hopkins Press, 1923.

Sethi, Rajiv and Muhamet Yildiz, "Public Disagreement," February 2009. Unpublished.

Sinclair, Andrew, *Prohibition: The Era of Excess*, London: Faber and Faber.

Smith, Alfred, *The New York Red Book*, Vol. 1923, Albany, NY: J. B. Lyon Co., 1923.

Spector, David, "Rational Debate and One-Dimensional Conflict," *Quarterly Journal of Economics*, February 2000, *115* (1), 181–200.

Stayton, W. H., "Our Experiment in National Prohibition. What Progress Has it Made?," *Annals of the American Academy of Political and Social Science*, 1923, *109*, 26–38.

Strulovici, Bruno, "Learning While Voting; Determinants of Collective Experimentation," *Econometrica*, 2010, *78* (3), 933–971.

Szymansky, Anne-Marie, *Pathways to Prohibition*, Durham, NC: Duke University Press, 2003.

the Amendment, Association Against Prohibition, *Scandals of Prohibition Enforcement*, Washington D.C: National Press Building, 1929.

the Prohibition Amendment, Association Against, *Measuring the Liquor Tide*, Washington D.C: National Press Building, 1930.

Tomkins, Floyd, "Prohibition," *Annals of the American Academy of Political and Social Science*, 1923, *109*, 15–25.

Tydings, Millard E., *Before and After Prohibition*, New York: The Macmillan Company, 1930.

Vives, Xavier, *Information and Learning in Markets: The Impact of Market Microstructure*, Princeton, NJ: Princeton University Press, 2010.

Wickersham-Commission, "Wickersham Commission Records, 1928-1931," Original Papers and Documents, Harvard University Law School Library 1928-1931.

_ , *Report on the Cost of Crime*, Vol. 5, United States Government Printing Office, 1931.

Wilcox, Walter, "An Attempt to Measure Public Opinion About Repealing the Eighteenth Amendment," *Journal of the American Statistical Association*, 1931, *26* (175), 243–261.

Wooldrige, Jeffrey, *Econometric Analysis of Cross Section and Panel Data*, MIT Press, 2002.

_ , "Simple Solutions to the Initial Conditions Problem in Dynamic, Nonlinear Panel Data Models with Unobserved Heterogeneity," *Journal of Applied Econometrics*, 2005, *20* (1), 39–54.

Tables

Table 1: Prohibition Enforcement during the 1920s

	Federal and State-Level Prohibition Enforcement	1923-1924	1925-1926	1927-1928	1929-1930	1931-1932
Midwest	Seized Distilleries, Stills and Still Worms	8,098	9,171	10,207	11,369	5,095
	Seized Fermenters	6,830	7,525	48,748	81,178	8,714
	Seized Spirits*	1,050,017	610,440	472,922	475,840	505,713
	Seized Malt, Wine, Cider, Mash and Pomace*	3,675,499	7,034,847	11,206,588	15,029,002	15,915,534
	Seized Autos and Boats	1,095	1,873	2,949	3,069	4,154
	Killed or Injured Officers	14	15	38	48	4
	Federal Arrests	-	24,150	28,185	31,755	25,528
	State Arrests	-	8,335	7,500	6,227	7,460
Northeast	Seized Distilleries, Stills and Still Worms	4,191	3,456	11,136	6,960	4,511
	Seized Fermenters	2,506	3,411	50,079	53,973	9,264
	Seized Spirits*	206,411	767,086	1,142,467	929,877	1,753,629
	Seized Malt, Wine, Cider, Mash and Pomace*	1,143,955	6,334,026	11,811,643	13,930,121	21,233,629
	Seized Autos and Boats	1,082	4,078	4,333	3,334	4,115
	Killed or Injured Officers	4	24	26	40	3
	Federal Arrests	-	35,316	46,396	29,657	47,585
	State Arrests	-	3,610	3,828	3,708	2,848
South	Seized Distilleries, Stills and Still Worms	34,087	48,038	43,327	43,273	31,671
	Seized Fermenters	179,280	238,528	267,605	301,521	8,513
	Seized Spirits*	284,888	732,713	647,875	799,950	842,375
	Seized Malt, Wine, Cider, Mash and Pomace*	11,933,042	22,026,530	30,428,757	34,753,824	29,953,542
	Seized Autos and Boats	2,703	4,979	5,290	7,058	9,054
	Killed or Injured Officers	18	42	62	122	4
	Federal Arrests	-	42,673	49,498	53,300	51,976
	State Arrests	-	11,778	11,001	12,811	7,737
West	Seized Distilleries, Stills and Still Worms	4,691	7,089	6,334	4,956	2,141
	Seized Fermenters	4,855	14,463	21,291	18,711	1,745
	Seized Spirits*	115,377	230,132	228,029	269,567	283,214
	Seized Malt, Wine, Cider, Mash and Pomace*	1,171,349	3,019,286	3,952,164	7,374,756	5,242,851
	Seized Autos and Boats	1,138	1,421	1,774	2,040	3,009
	Killed or Injured Officers	18	20	31	17	3
	Federal Arrests	-	17,007	14,486	16,845	9,943
	State Arrests	-	4,718	6,828	6,640	9,262

*Gallons
Northeast includes ME, NH, VT, MA RI, CT, NY, PA, and NJ
Midwest includes ND, SD, NE, KS, MN, IA, MO, WI, IL, IN, MI, OH
South includes DE, MD, DC, VA, WV, KY, NC, TN, SC, GA, AL, MS, FL, AR, LA, OK, and TX
West includes WA, OR, CA, ID, MT, WY, CO, UT, NV, AZ, NM
Source: U.S. Bureau of Prohibition, Statistics Concerning Intoxicating Liquors, and Wickersham Commission papers.

Table 2: Summary Statistics

Summary Statistics	Region*			
	Midwest	Northeast	South	West
Dry Religions				
% Baptist (1916)	12.066	8.601	29.660	16.149
	(8.559)	(5.901)	(13.921)	(9.245)
% Evangelical (1916)	2.623	0.647	0.986	0.852
	(2.083)	(1.393)	(1.496)	(0.510)
% Mormon (1916)	0.298	0.057	0.160	5.198
	(0.578)	(0.140)	(0.413)	(18.856)
% Methodist/Episcopal (1916)	11.666	10.269	24.825	13.740
	(6.356)	(4.740)	(9.855)	(6.469)
% Presbyterian (1916)	4.779	4.471	6.060	8.068
	(2.642)	(3.521)	(3.246)	(4.743)
Wet Religions				
% Eastern Orthodox (1916)	1.139	1.272	0.437	1.890
	(1.443)	(1.461)	(0.591)	(2.222)
% Jewish (1916)	1.741	2.771	1.716	1.887
	(1.130)	(2.515)	(1.303)	(1.033)
% Lutheran (1916)	7.677	3.510	2.521	3.181
	(5.708)	(3.678)	(2.119)	(1.944)
% Catholic (1916)	55.798	66.631	31.496	47.494
	(16.245)	(12.568)	(21.798)	(20.723)
Demographics				
% Black (1910)	3.104	2.325	26.525	1.336
	(2.823)	(2.057)	(12.398)	(0.951)
% Foreign White (1910)	24.573	32.618	7.928	23.083
	(10.557)	(8.664)	(5.705)	(5.364)
% Native White (1910)	72.220	64.945	65.412	73.210
	(9.036)	(8.170)	(9.816)	(5.610)
% Ages 15-24 (1910)	20.840	20.041	21.050	18.970
	(0.909)	(1.010)	(1.166)	(0.909)
% Ages 25-44 (1910)	34.258	33.515	34.159	38.519
	(1.610)	(1.204)	(2.163)	(3.087)
Legislation				
Number of Dry Laws (1919)	6.709	3.376	5.999	9.303
	(3.136)	(2.014)	(1.696)	(3.310)
Number of Years Under Prohibition**	15.925	15.018	18.035	15.844
	(4.443)	(0.169)	(4.080)	(1.253)

Outcomes	Midwest 1910s	Midwest 1920s	Northeast 1910s	Northeast 1920s	South 1910s	South 1920s	West 1910s	West 1920s
Per Capita Police Expenditure (1913 prices)	1.541	1.793	1.977	2.312	1.376	1.716	1.511	1.796
	(0.726)	(0.759)	(0.732)	(0.798)	(0.513)	(0.800)	(0.687)	(0.699)
Police Expenditure Share	0.108	0.092	0.112	0.096	0.123	0.112	0.087	0.081
	(0.032)	(0.026)	(0.021)	(0.020)	(0.024)	(0.026)	(0.026)	(0.017)
Drunkenness Arrest Rate (per 1,000)	16.000	14.459	16.653	12.132	18.273	18.606	22.560	13.963
	(11.191)	(8.591)	(18.875)	(13.081)	(11.525)	(10.786)	(14.801)	(6.241)
Homicide Rate (per 100,000)	10.807	18.124	5.368	10.076	22.849	28.132	9.897	11.620
	(5.730)	(7.560)	(1.776)	(3.633)	(18.085)	(17.150)	(3.375)	(3.379)
% Anti-Prohibition vote share***	0.518	0.723	0.523	0.826	0.467	0.577	0.457	0.678
	(0.169)	(0.168)	(0.103)	(0.129)	(0.166)	(0.220)	(0.138)	(0.116)

*Regions as as classified by the Bureau of the Census: North East includes ME, NH, VT, MA, RI, CT, NY, PA, and NJ
Midwest includes ND, SD, NE, KS, MN, IA, MO, WI, IL, IN, MI, OH
South includes DE, MD, DC, VA, WV, KY, NC, TN, SC, GA, AL, MS, FL, AR, LA, OK, and TX
West includes WA, OR, CA, ID, MT, WY, CO, UT, NV, AZ, and NM
** During the 1910-1933 period
***From state level referenda
Standard Deviations in parenthesis
All summary statistics are weighted by city population

Table 3: Public Opinion Regressions

Electoral Support for Prohibition

Panel A															
Dependent Variable	Wet Vote Share														
Sample	County Sample										City Sample				
	All Counties					Counties with Pop>30,000									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)
Post-Prohibition Indicator	0.1373	-0.1099	0.0806	-0.0572	-0.1088	0.1714	-0.2613	-0.0485	-0.1145	-0.2505	0.2453	-0.0779	-0.1307	0.1601	0.0215
	(0.004)	(0.018)	(0.022)	(0.035)	(0.017)	(0.009)	(0.049)	(0.062)	(0.054)	(0.051)	(0.013)	(0.084)	(0.123)	(0.097)	(0.094)
"Wetness"		0.1201	0.1422	0.1211	0.1073		0.4879	0.5443	0.2058	0.4484		0.2221	-0.2573	-0.7241	-0.0004
		(0.090)	(0.095)	(0.078)	(0.096)		(0.310)	(0.309)	(0.241)	(0.308)		(0.363)	(0.370)	(0.495)	(0.413)
Baseline "Wetness" x Post-Prohibition Indicator		0.6969	0.5965	0.4808	0.7349		1.0874	0.8798	0.5082	1.1367		0.6643	0.5896	0.4258	0.4740
		(0.045)	(0.042)	(0.044)	(0.047)		(0.118)	(0.121)	(0.112)	(0.129)		(0.181)	(0.158)	(0.125)	(0.213)
log of Population			0.0881	0.0671	0.0806			0.0821	0.0751	0.0392			0.1606	0.0935	0.0111
			(0.011)	(0.008)	(0.011)			(0.027)	(0.022)	(0.026)			(0.058)	(0.062)	(0.048)
Urban share of county			0.0213	0.0517	0.0572			-0.0815	0.0588	0.0262			-0.1834	-0.1455	-0.0921
			(0.039)	(0.028)	(0.041)			(0.127)	(0.086)	(0.128)			(0.082)	(0.120)	(0.099)
Number of Dry Laws													-0.0342	-0.0019	-0.0222
													(0.011)	(0.035)	(0.008)
Referendum Year			-0.0109					-0.0102					0.0005		
			(0.001)					(0.002)					(0.005)		
Referendum Type:															
Prohibition Law			-0.0048					0.0072					0.0024		
			(0.008)					(0.014)					(0.028)		
Constitutional Convention Election			0.0100					0.0386					0.1141		
			(0.010)					(0.016)					(0.045)		
Inverse Mills Ratio x Pre-prohibition Period					0.1577					0.1896					-0.0546
					(0.031)					(0.055)					(0.075)
State cross Post-Prohibition Effects	No	No	No	Yes	No	No	No	No	Yes	No	No	No	No	Yes	No
Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R squared	0.385	0.488	0.555	0.766	0.525	0.498	0.646	0.688	0.876	0.665	0.741	0.774	0.813	0.943	0.792
No. of Cross Sections	1693	1693	1693	1693	1693	672	672	672	672	672	258	258	258	258	258
No. of Observations	3386	3386	3386	3386	3386	337	337	337	337	337	129	129	129	129	129

Panel B															
	Probit Selection Equation														
Share of Wet Religions					-1.092					-1.092					-1.092
					(1.277)					(1.277)					(1.277)
Share of Non Native White					-7.297					-7.297					-7.297
					(3.836)					(3.836)					(3.836)
Constant					3.630					3.630					3.630
					(1.605)					(1.605)					(1.605)
Pseudo R squared					0.173					0.173					0.173
No. of Observations					31					31					31
Log Likelihood					-11.32					-11.32					-11.32

Note: Constant for the Vote share equations not reported. Standard Errors are robust and clustered at the county or city level.
The Selection equation is a probit at the state level, of an Indicator for having had an Alcohol referendum in the Pre-Prohibition period,
on the state-level share of adherents to any "wet" religion (Orthodox, Jewish, Luthera, Catholic, Other) from the 1916 Census of Religions,
and the share of non-native white individuals in the population, from the 1910 Population Census.

Table 4: Conditional Maximum Likelihood Estimates

Conditional Maximum Likelihood Estimates

	Covariate	Coefficients		Covariate	Coefficients		Covariate	Coefficients
a	% ages 15-44*	0.606	α	Constant	2.758	**B**	% Baptist in 1911	-0.237
		(0.273)			(0.762)			(0.071)
	% Foreign White*	0.519		Number of Alcohol-Related Laws	0.117		% Orthodox in 1911	0.022
		(0.198)			(0.248)			(0.020)
	% Black*	0.199		Enforcement Law	-0.895		% Evangelical in 1911	-0.034
		(0.058)			(0.206)			(0.008)
	% Baptist*	0.382		Prohibition Unit Seat	-0.227		% Jewish in 1911	-0.021
		(0.281)			(0.153)			(0.031)
	% Orthodox*	0.434		Prohibition Unit: Providence	-0.281		% Mormon in 1911	-0.019
		(0.223)			(0.201)			(0.677)
	% Evangelical*	0.589		Prohibition Unit: Washington	-0.909		% Lutheran in 1911	-0.116
		(0.498)			(0.496)			(0.078)
	% Jewish*	-0.447		Prohibition Unit: Jacksonville	-1.090		% Methodist in 1911	-0.289
		(0.462)			(1.432)			(0.161)
	% Mormon*	0.257		Prohibition Unit: Detroit	-0.851		% Catholic in 1911	-0.738
		(0.259)			(0.326)			(0.619)
	% Lutheran*	0.767		Prohibition Unit: Chicago	-2.065		% Presbyterian in 1911	-0.034
		(0.295)			(1.392)			(0.011)
	% Methodist*	0.746		Prohibition Unit: Kansas City	-1.166		Constant	-2.231
		(0.555)			(0.725)			(0.612)
	% Catholic*	0.174		Prohibition Unit: San Francisco	2.125	σ_q^2		0.277
		(0.041)			(1.217)			(0.084)
	% Presbyterian*	-0.255		Prohibition Unit: Los Angeles	0.004	σ_z^2		0.061
		(0.016)			(0.603)			(0.028)
	Log of Population*	0.001		Prohibition Unit: Seattle	1.669	σ_ξ^2		0.348
		(0.034)			(0.659)			(0.135)
	Constant	2.477	**z**	% Baptist	0.890	σ_ζ^2		1.000
		(0.088)			(0.395)			(0.000)
b		8.666		% Orthodox	-0.223	ρ		-0.495
		(2.773)			(0.076)			(0.089)
χ		8.948		% Evangelical	0.187			
		(3.454)			(0.082)			
κ		0.266		% Jewish	0.728			
		(0.071)			(0.650)			
λ		0.259		% Mormon	0.172			
		(0.108)			(0.212)			
θ	Border	-0.302		% Lutheran	0.071			
		(0.101)			(0.296)			
	South	-0.652		% Methodist	1.186			
		(0.193)			(0.438)			
	State Capital	-0.248		% Catholic	0.819			
		(0.173)			(0.669)			
	Share Ages 15-44*	0.188		% Presbyterian	-0.772			
		(0.498)			(1.370)			
	Share Foreign White*	0.073		Constant	0.024			
		(0.059)			(0.015)			
	Share Black*	0.231						
		(0.088)						
	Constant	1.504						
		(0.713)						
	Log-likelihood:	6560.774						
	Observations:	1254						

*1911-1929 averages

Note: Standard Errors are computed through a bootstrap of size 100.

Estimates of the State Effects are omitted from the table.

Table 5: Mean Parameter estimates

Average Estimated Values of Parameter Estimates	
Model Parameters	Mean Estimate
a	3.167
	(0.055)
b	8.666
	(0.000)
χ	8.948
	(0.000)
κ	0.266
	(0.000)
λ	0.259
	(0.000)
α	3.437
	(1.010)
z	0.687
	(0.108)
θ	1.377
	(0.258)
B	-1.311
	(0.265)

Note: Standard Deviations in parenthesis

Figures

Figure 1: The Homicide Rate in U.S. Cities, 1911-1936 (per 100,000)[42]

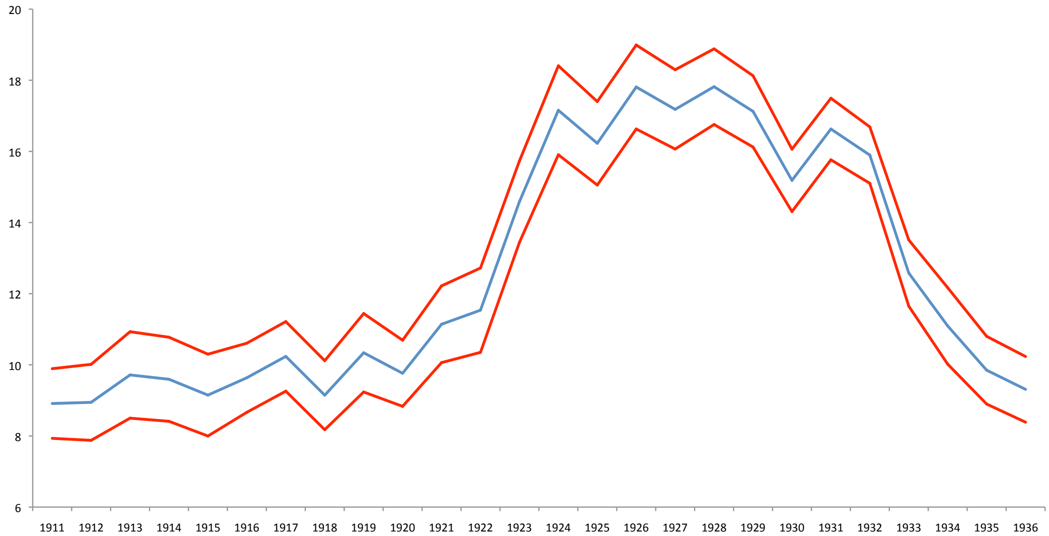
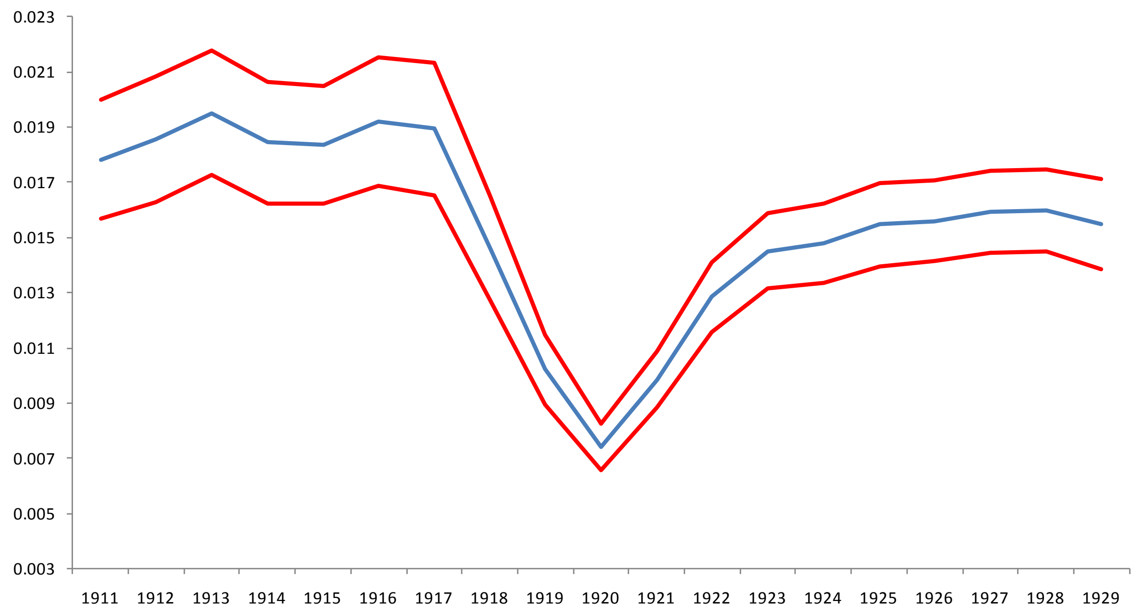


Figure 2: The Drunkenness Arrest Rate in U.S. Cities, 1911-1929 (per capita)



[42]In all figures, red lines are 95% confidence intervals.

Figure 3: Timing of State Adoption of Prohibition

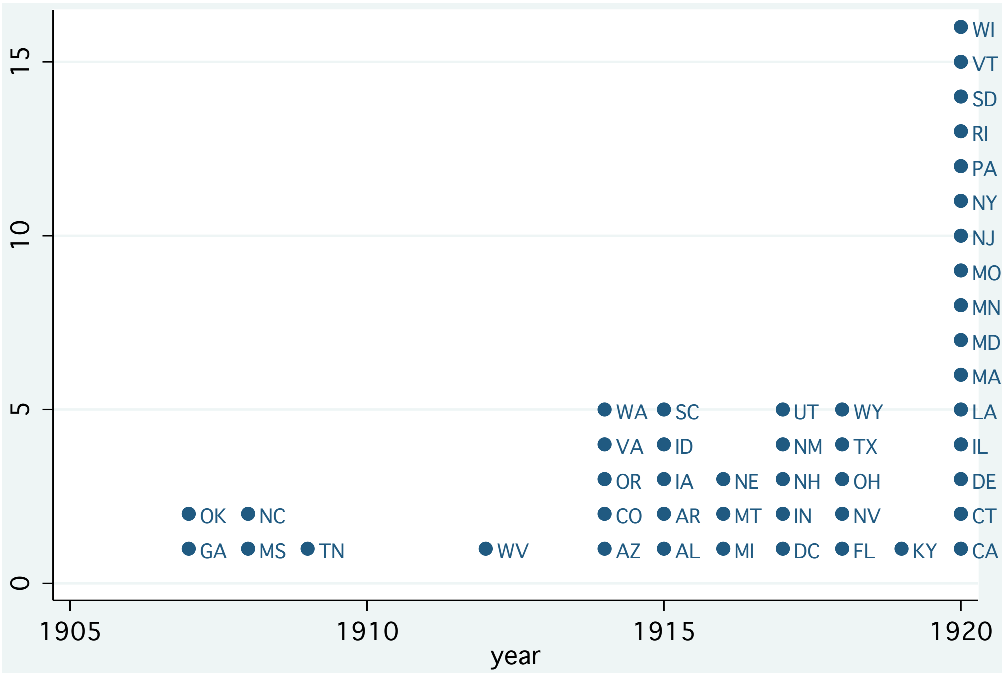
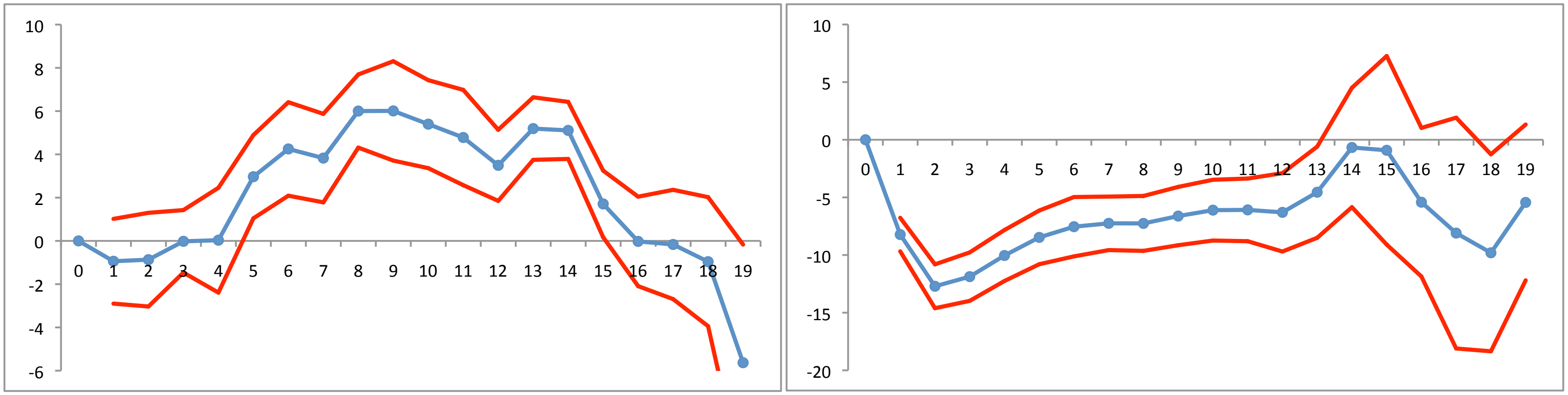


Figure 4: δ_τ's from equation (1):

Panel A: Homicide Rate (per $100,000$) (left) and Drunkenness Arrest Rate (per $1,000$) (right)



Panel B: Police Expenditure Share (left) and Per Capita Police Expenditure (right)

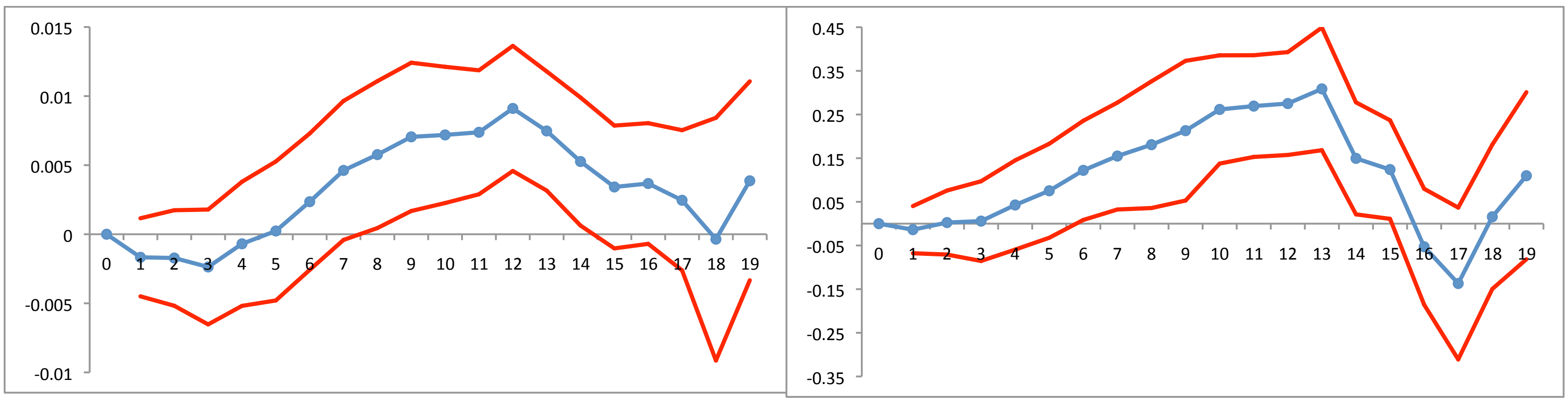
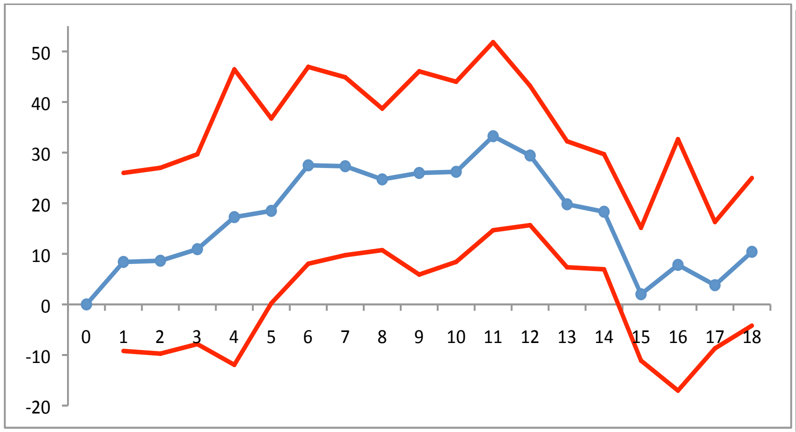
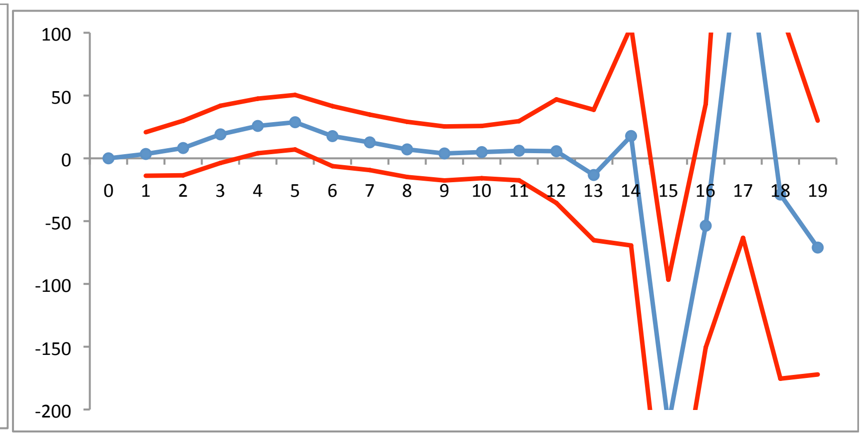


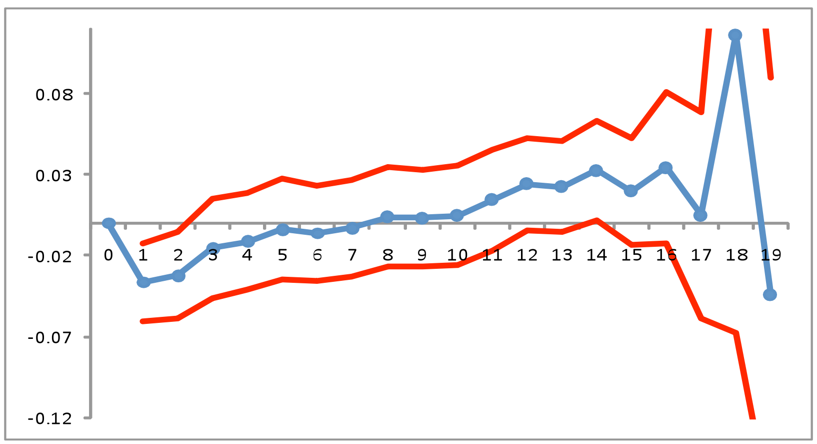
Figure 5: ϕ_τ's from Equation (3):

Panel A: Homicide Rate (per 100,000) (left) and Drunkenness Arrest Rate (per 1,000) (right)





Panel B: Police Expenditure Share (left) and Per Capita Police Expenditure (right)



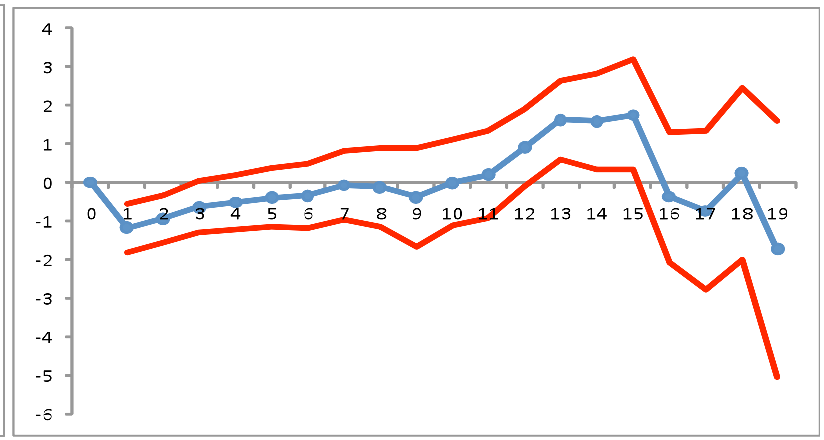


Figure 6: Alcohol Referenda and Public Opinion shift

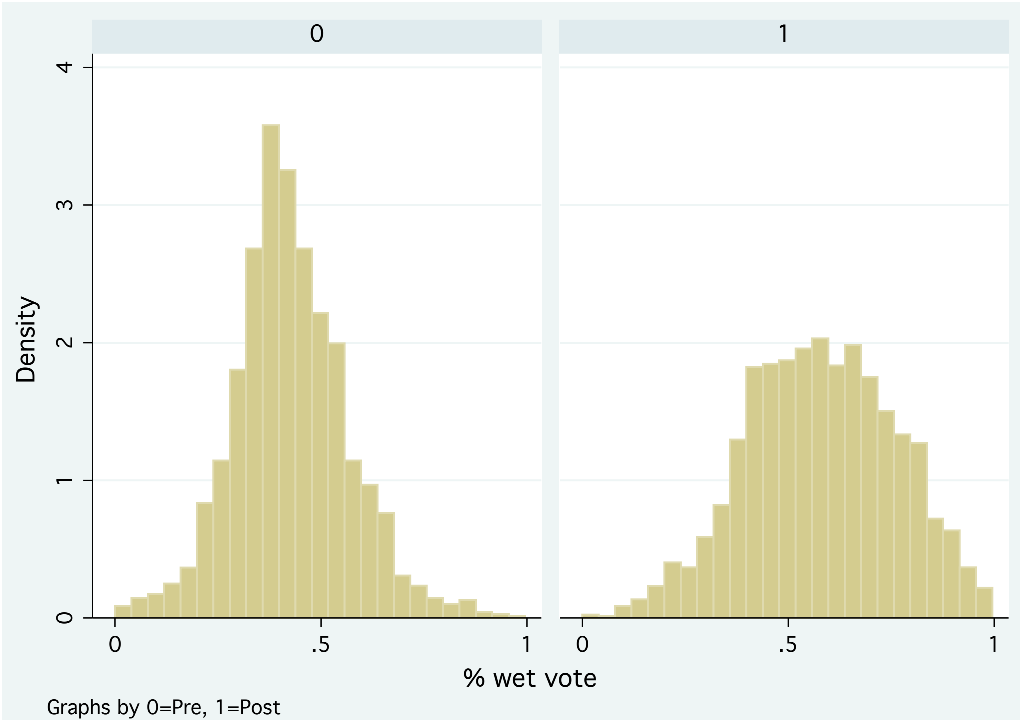


Figure 7: Moral Views and Changes in Public Opinion (U.S. counties)

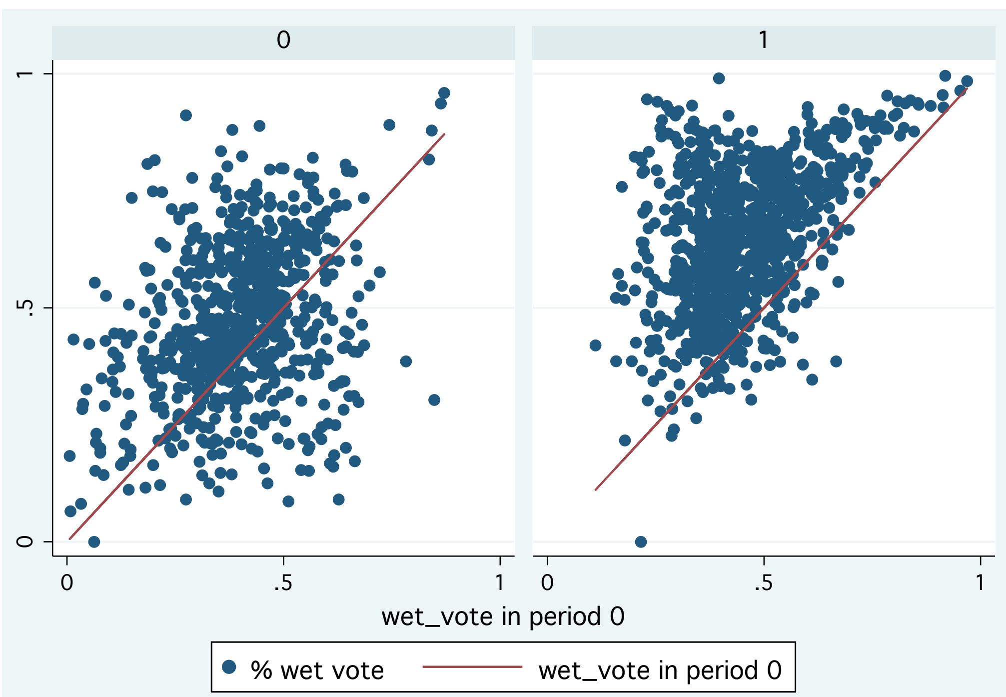
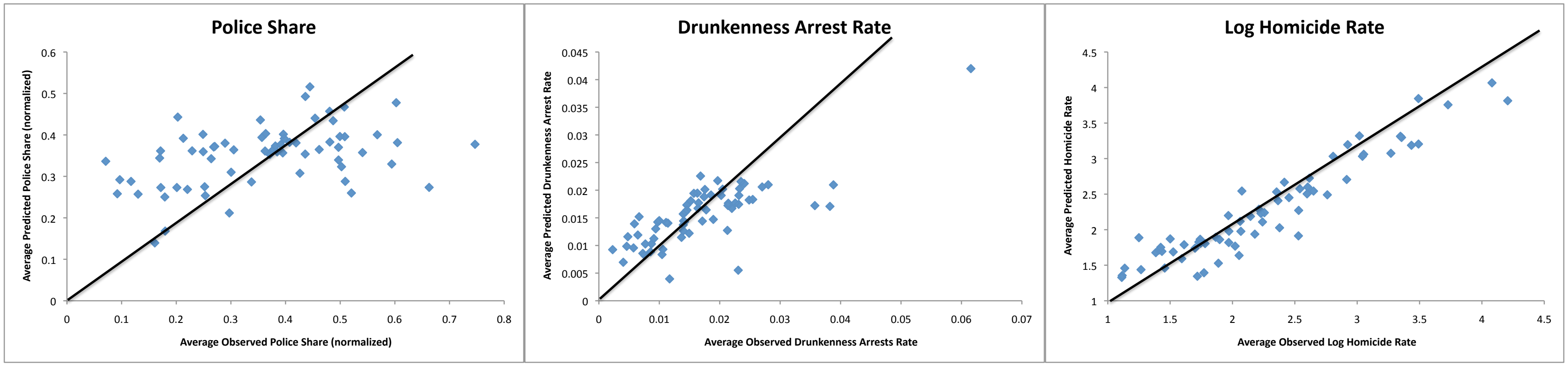


Figure 8: Fit of the Model

Panel A: Cross-Sectional Fit



Panel B: Time-Series Fit

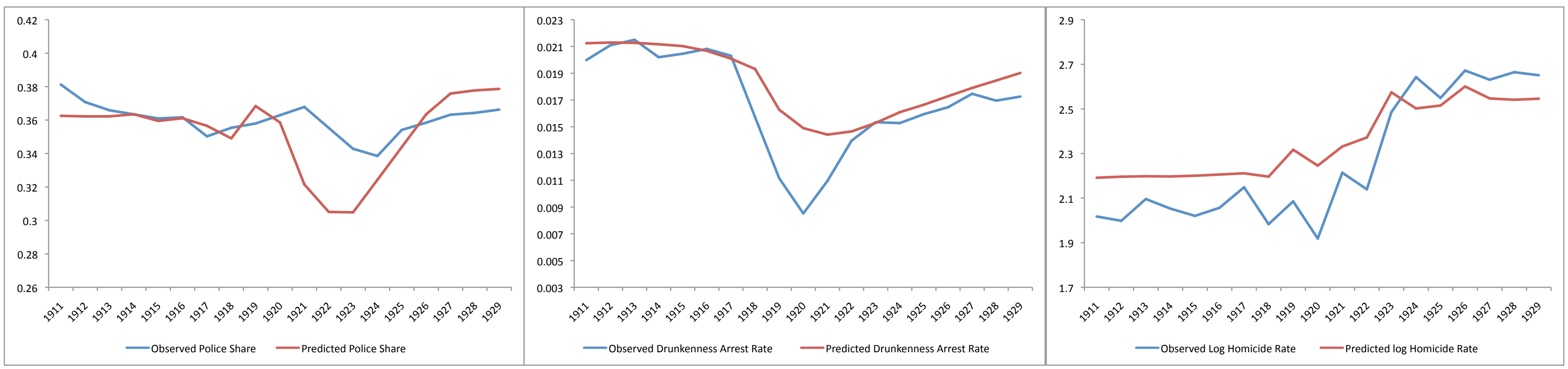


Figure 9: Distribution of the Estimated θ_c's

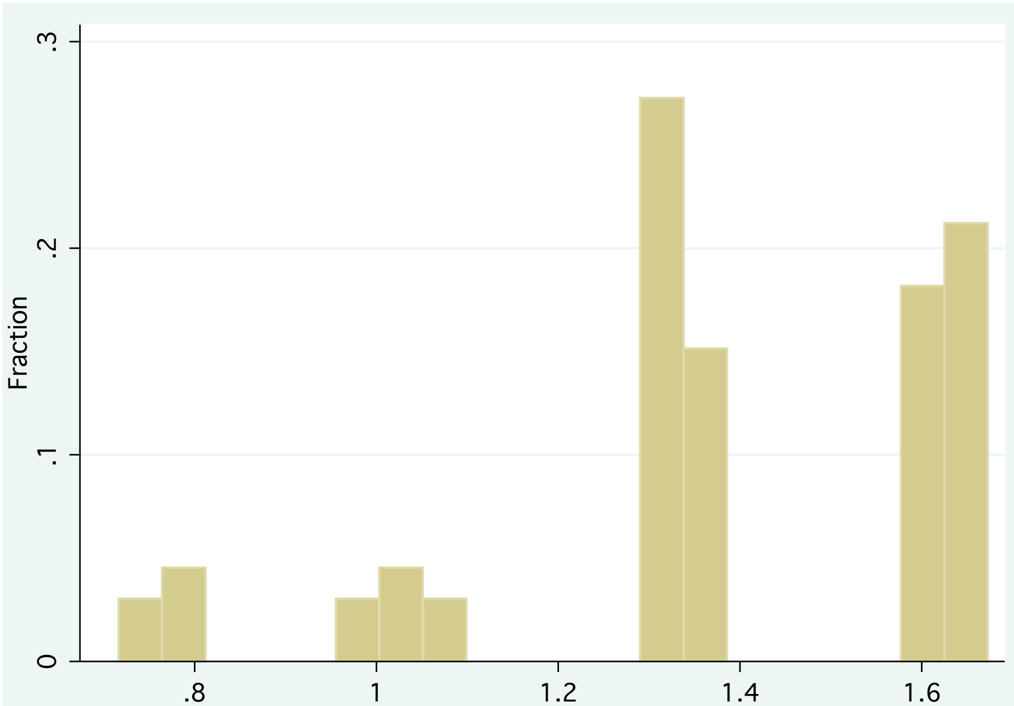


Figure 10: Estimated Prior Beliefs B_c vs. Estimated Average Moral views $\overline{z}_c$.

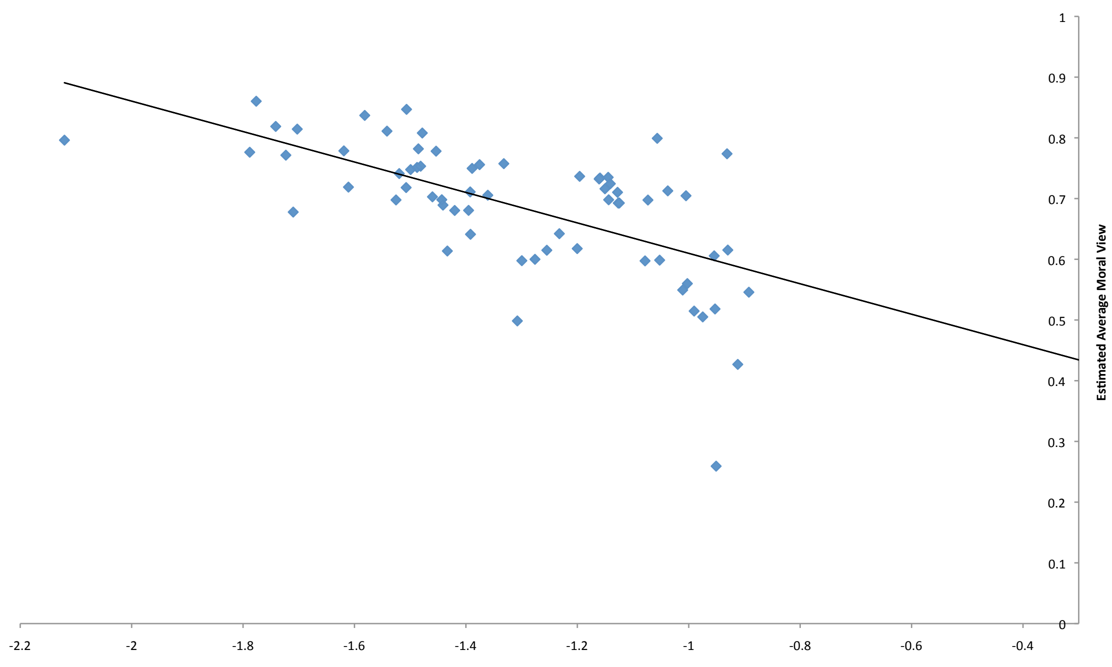


Figure 11: Estimated Functional Forms

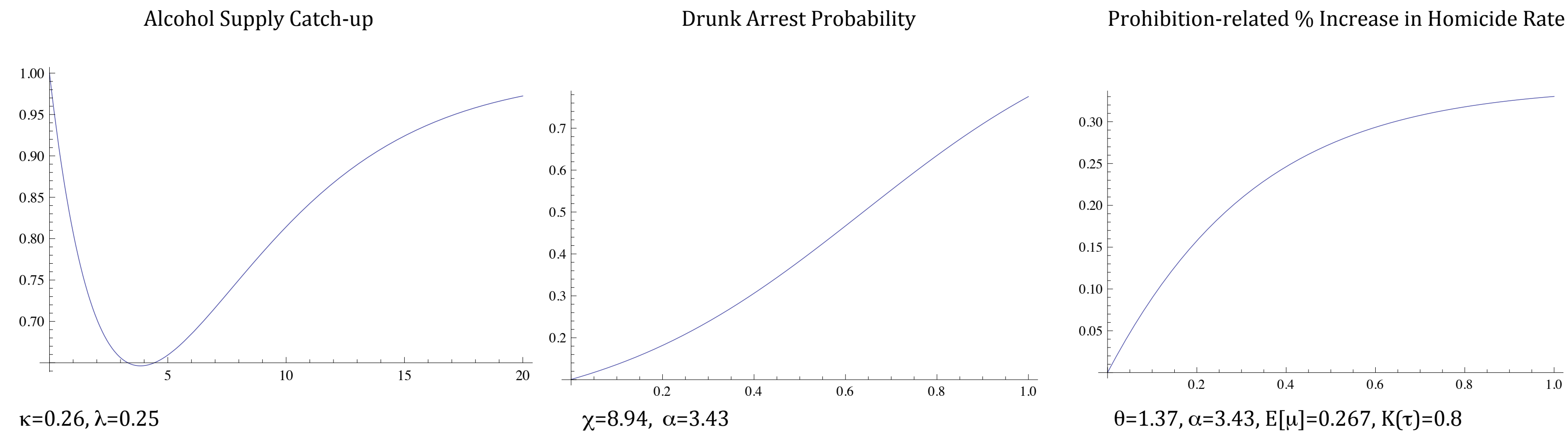


Figure 12: Estimated Densities of the Median Voters' Unobserved ϱ^{med}

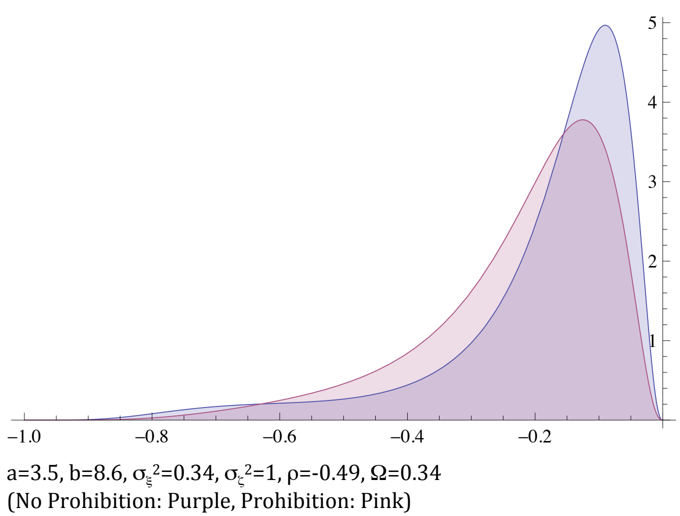


Figure 13: Estimated Belief Sequences: Empirical Distribution

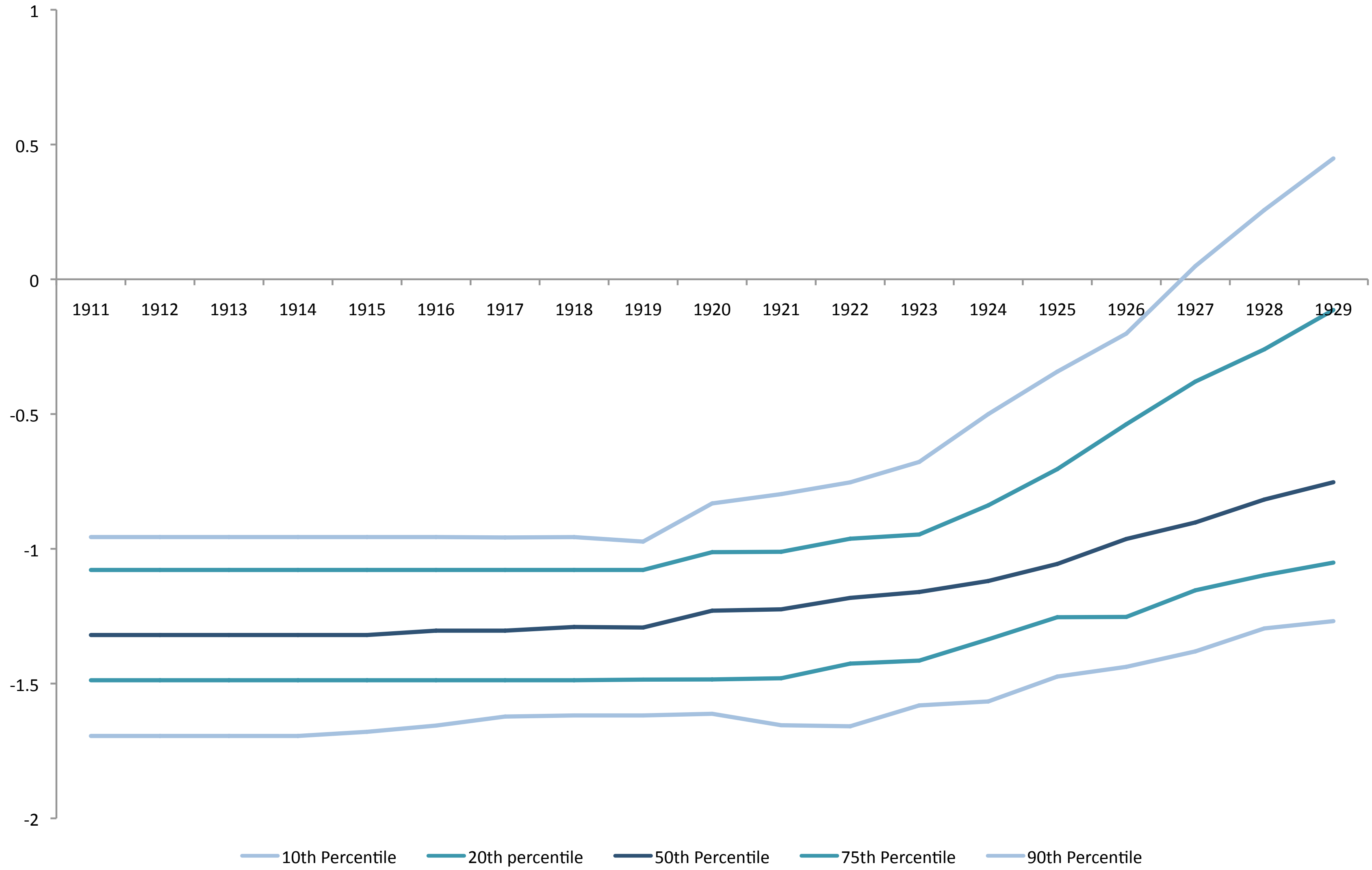
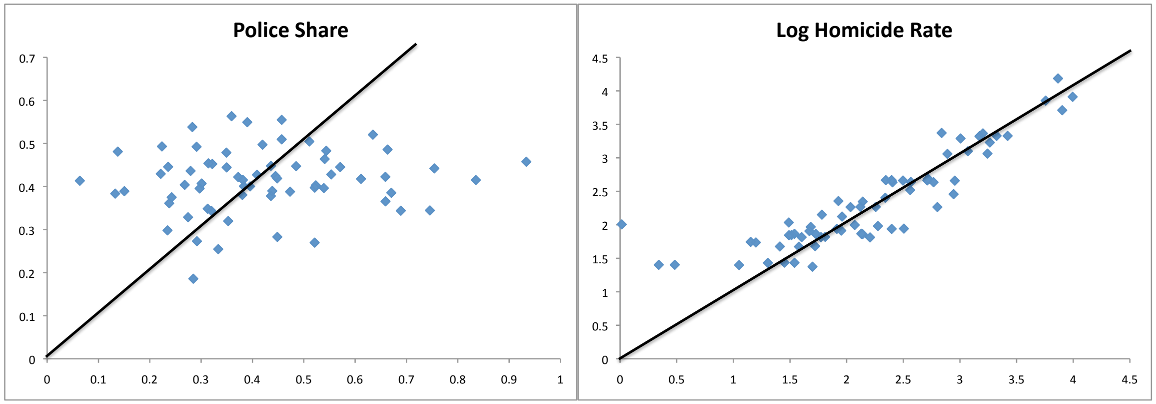
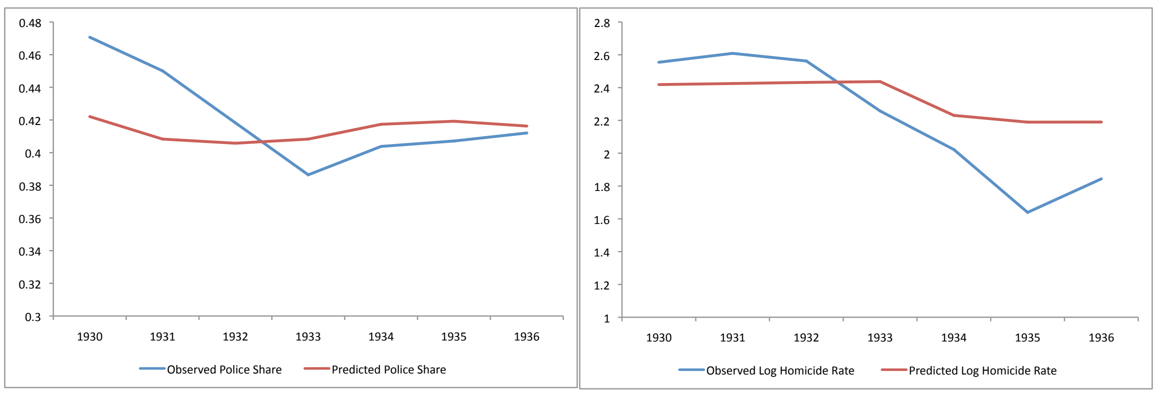


Figure 14: “Out of Sample” predictions for the years 1930-1936

Panel A: Cross-Sectional Fit



Panel B: Time-Series Fit



Appendix 1: Derivation of Ideal Law Enforcement Choice

Indirect utility under no Prohibition is given by

$$E_tU^i(p_t|P_t=0) = 1_{\{i\in W_t\}}exp(-\alpha_t p_t) - z^i\frac{a}{a+b}exp(-\alpha_t p_t) + exp(1-p_t) - \Theta_S - \frac{a}{a+b}exp(-\alpha_t p_t) \quad (26)$$

The first order condition with respect to p_{ct} from equation (26) is

$$-1_{\{i\in W_{ct}\}}\alpha_{ct}exp(-\alpha_{ct}p_{ct}) + \alpha_{ct}\frac{a_c}{a_c+b}z_c^i exp(-\alpha_{ct}p_{ct}) - exp(1-p_{ct}) + \alpha_{ct}\frac{a_c}{a_c+b}exp(-\alpha_{ct}p_{ct}) \leq 0$$

Solving for p_{ct}, equation (16) directly follows. The second order condition for this problem is given by

$$1_{\{i\in W_{ct}\}}\alpha_{ct}^2 exp(-\alpha_{ct}p_{ct}) - \alpha_{ct}^2\frac{a_c}{a_c+b}z_c^i exp(-\alpha_{ct}p_{ct}) + exp(1-p_{ct}) - \alpha_{ct}^2\frac{a_c}{a_c+b}exp(-\alpha_{ct}p_{ct}) < 0$$

$$\Leftrightarrow 2ln\alpha_{ct} + ln\left[\frac{a_c}{a_c+b}(z_c^i+1) - 1_{\{i\in W_t\}}\right] - 1 > (\alpha_{ct}-1)p_{ct} \quad (27)$$

I verify this condition is satisfied for the parameter estimates.

Indirect utility under Prohibition is given by

$$E_tU^i(p_t|P_t=1) = 1_{\{i\in W_t\}}k(\tau_t)exp(-\alpha_t p_t) - z^i\frac{a}{a+b}k(\tau_t)exp(-\alpha_t p_t)$$

$$+ exp(1-p_t) - \Theta_S - \frac{a}{a+b}k(\tau_t)exp(-\alpha_t p_t) - \overline{\theta}_t^i k(\tau_t)\frac{a}{a+b}[1-exp(-\alpha_t p_t)] \quad (28)$$

The first order condition with respect to p_{ct} from equation (28) is

$$-1_{\{i\in W_{ct}\}}k(\tau_{ct})\alpha_{ct}exp(-\alpha_{ct}p_{ct}) + \alpha_{ct}\frac{a_c}{a_c+b}k(\tau_{ct})z_c^i exp(-\alpha_{ct}p_{ct}) - exp(1-p_{ct})$$

$$+\alpha_{ct}\frac{a_c}{a_c+b}k(\tau_{ct})exp(-\alpha_{ct}p_{ct}) - \alpha_{ct}\overline{\theta}_{ct}^i\frac{a_c}{a_c+b}k(\tau_{ct})exp(-\alpha_{ct}p_{ct}) \leq 0$$

Solving for p_{ct}, equation (17) directly follows. The second-order condition for the solution in equation (17) to be a maximum is

$$\Leftrightarrow 2ln\alpha_{ct} + ln\left[\frac{a_c}{a_c+b}(z_c^i - \overline{\theta}_{ct}^i + 1) - 1_{\{i\in W_t\}}\right] - 1 > (\alpha_{ct}-1)p_{ct} \quad (29)$$

Appendix 2: Proof of Proposition 1

Proof. In this community there are three sources of heterogeneity in preferences over law enforcement: the distribution of moral views, the distribution of belief biases, and the distribution of types (wet and dry). First, observe that conditional on (ζ^i, ξ^i), the preferred level of law enforcement of a wet voter is shifted down by a constant factor relative to the optimal choice of a dry individual. Thus, for periods under Prohibition define $\varrho^i_{DP} \equiv \frac{a}{a+b}(\zeta^i - \Omega_t \frac{1}{\sigma^2_\xi}\xi^i)$ (DP for Dry under Prohibition), and $\varrho^i_{WP} \equiv \frac{a}{a+b}(\zeta^i - \Omega_t \frac{1}{\sigma^2_\xi}\xi^i) - 1$ (WP for Wet under Prohibition). These are normal random variables distributed according to $\varrho^i_{DP} \sim N(0, \sigma^2_{\varrho Pt})$ and $\varrho^i_{WP} \sim N(-1, \sigma^2_{\varrho Pt})$ respectively, where $\sigma^2_{\varrho Pt} \equiv \left(\frac{a}{a+b}\right)^2 \left(\sigma^2_\zeta + \Omega^2_t \frac{1}{\sigma^2_\xi} - 2\Omega_t \rho \frac{\sigma_\zeta}{\sigma_\xi}\right)$[43]. Now define $\varrho^i_P \equiv 1_{\{i \in D_t\}}\varrho^i_{DP} + 1_{\{i \in W_t\}}\varrho^i_{WP}$. The conditional density of ϱ^i_P is given by

$$f_{\varrho P}(\varrho^i_P|\mu_t) = (1-\mu_t)N(0, \sigma^2_{\varrho Pt}) + \mu_t N(-1, \sigma^2_{\varrho Pt})$$

since with probability μ_t a wet individual is sampled, and with probability $1-\mu_t$ a dry individual is sampled. Thus, the distribution of ϱ^i_P in the population is a mixture of two normal random variables with a common variance, one of which is shifted to the left by 1 relative to the other. Given the normality of ϱ^i_{WP} and ϱ^i_{DP}, as $\mu_t \to 0$, the median voter's type $\varrho^{med}_P \to 0$, and as $\mu_t \to 1$, $\varrho^{med}_P \to -1$, so that $\varrho^{med}_P \in (-1, 0)$. For periods under no Prohibition, analogously define $\varrho^i_{DN} \equiv \frac{a}{a+b}\zeta^i$ (DN for dry under no Prohibition) and $\varrho^i_{WN} \equiv \frac{a}{a+b}\zeta^i - 1$ (WN for wet under no Prohibition), which are distributed according to $\varrho^i_{DN} \sim N(0, \sigma^2_{\varrho N})$ and $\varrho^i_{WN} \sim N(-1, \sigma^2_{\varrho N})$ respectively, with $\sigma^2_{\varrho N} \equiv \left(\frac{a}{a+b}\right)^2 \sigma^2_\zeta$. Now define $\varrho^i_N \equiv 1_{\{i \in D_t\}}\varrho^i_{DN} + 1_{\{i \in W_t\}}\varrho^i_{WN}$, which is a random variable whose conditional density is given by

$$f_{\varrho N}(\varrho^i_N|\mu_t) = (1-\mu_t)N(0, \sigma^2_{\varrho N}) + \mu_t N(-1, \sigma^2_{\varrho N})$$

Indirect preferences over law enforcement in (16) and (17) can be expressed in terms of ϱ^i_N and ϱ^i_P. It follows that this is a purely private-values election because individuals realize that differences in beliefs are due to individual-specific biases. For a given individual, the voting decisions of the members of his community do not convey any additional information. Moreover, indirect preferences over law enforcement are single-peaked in ϱ^i_j, so the Median Voter Theorem holds, and the unique political equilibrium value of p_t is given by the preferred choice of law enforcement of the median over the distribution of ϱ^i_j, conditional on μ_t.

The (conditional) median for Prohibition years will be given by the value of ϱ^{med}_P which solves the following equation

$$(1-\mu_t)\int_{-\infty}^{\varrho^{med}_P(\mu_t)} \frac{1}{\sqrt{2\pi}\sigma_{\varrho P}} exp\left(-\frac{1}{2\sigma^2_{\varrho P}}\varrho^2\right) d\varrho + \mu_t \int_{-\infty}^{\varrho^{med}_P(\mu_t)} \frac{1}{\sqrt{2\pi}\sigma_{\varrho P}} exp\left(-\frac{1}{2\sigma^2_{\varrho P}}(\varrho+1)^2\right) d\varrho = \frac{1}{2} \quad (30)$$

where I have made explicit the dependence of ϱ^{med}_P on the wet share in the community. Because the realization of μ_t is unobserved, the median ϱ^{med}_P in the population as defined in (30) is a random

[43] This variance is time-varying. As learning takes place and $\Omega_t \to 0$, $\sigma^2_{\varrho Pt} \to \sigma^2_\zeta$.

variable whose density is derived below. The equation analogous to (30) implicitly defining ϱ_N^{med} (the conditional median of the distribution of ϱ_N^i) and its corresponding density are found analogously[44].

Derivation of the density of ϱ_P^{med}:

First, recall that $f_\mu(\mu; a, b)$, the density of μ_{ct}, is beta with parameters (a_c, b). From (30), μ_{ct} can be directly expressed as a function of ϱ_P^{med}:

$$\mu_{ct} \equiv h_\mu(\varrho_P^{med}) = \frac{\frac{1}{2} - \Phi\left(\frac{\varrho_P^{med}}{\sigma_{\varrho Pt}}\right)}{\Phi\left(\frac{\varrho_P^{med}+1}{\sigma_{\varrho Pt}}\right) - \Phi\left(\frac{\varrho_P^{med}}{\sigma_{\varrho Pt}}\right)} \tag{31}$$

If this is a one-to-one mapping, the density of ϱ_P^{med} will be given by

$$f_{\varrho_P^{med}}(\varrho_P^{med}; a_c, b_c, \sigma_{\varrho Pt}) = f_\mu(h_\mu(\varrho_P^{med}); a_c, b_c) \left| \frac{\partial h_\mu(\varrho_P^{med})}{\partial \varrho_P^{med}} \right|$$

The derivative of h_μ is given by

$$\frac{\partial h_\mu(\varrho_P^{med})}{\partial \varrho_P^{med}} = \frac{\frac{1}{\sigma_{\varrho Pt}}\phi\left(\frac{\varrho_P^{med}}{\sigma_{\varrho Pt}}\right)\left[\frac{1}{2} - \Phi\left(\frac{\varrho_P^{med}+1}{\sigma_{\varrho Pt}}\right)\right] - \frac{1}{\sigma_{\varrho Pt}}\phi\left(\frac{\varrho_P^{med}+1}{\sigma_{\varrho Pt}}\right)\left[\frac{1}{2} - \Phi\left(\frac{\varrho_P^{med}}{\sigma_{\varrho Pt}}\right)\right]}{\left[\Phi\left(\frac{\varrho_P^{med}+1}{\sigma_{\varrho Pt}}\right) - \Phi\left(\frac{\varrho_P^{med}}{\sigma_{\varrho Pt}}\right)\right]^2} < 0$$

To see that $\frac{\partial h_\mu(\varrho_P^{med})}{\partial \varrho_P^{med}} < 0$ notice that the first term in square brackets is always smaller than the second term in square brackets. For $\varrho_P^{med} \geq 0$, $\phi\left(\frac{\varrho_P^{med}}{\sigma_{\varrho Pt}}\right) \geq \phi\left(\frac{\varrho_P^{med}+1}{\sigma_{\varrho Pt}}\right)$, and the first term in brackets is more negative than the second term in brackets, so the numerator is negative. For $\varrho_P^{med} < -\frac{1}{2}$, $\phi\left(\frac{\varrho_P^{med}}{\sigma_{\varrho Pt}}\right) < \phi\left(\frac{\varrho_P^{med}+1}{\sigma_{\varrho Pt}}\right)$, the second term in brackets is strictly positive, and the first term in brackets is also positive (but smaller than the second term in brackets), so the numerator is negative. For $\varrho_P^{med} \in (-\frac{1}{2}, 0)$, $\phi\left(\frac{\varrho_P^{med}}{\sigma_{\varrho Pt}}\right) \geq \phi\left(\frac{\varrho_P^{med}+1}{\sigma_{\varrho Pt}}\right)$, the first term in brackets is negative, and the second term in brackets is positive, so the numerator is negative.

Thus, h_μ is a one-to-one mapping, and the likelihood for ϱ_P^{med} is

$$f_{\varrho_P^{med}}(\varrho_P^{med}; a_c, b, \sigma_{\varrho Pt}) = \frac{1}{\sigma_{\varrho Pt}} \frac{1}{\int_0^1 v^{a_c-1}(1-v)^{b-1}dv} \times$$

$$\frac{\phi\left(\frac{\varrho^{med}+1}{\sigma_{\varrho Pt}}\right)\left[\frac{1}{2} - \Phi\left(\frac{\varrho^{med}}{\sigma_{\varrho Pt}}\right)\right]^{a_c}\left[\Phi\left(\frac{\varrho^{med}+1}{\sigma_{\varrho Pt}}\right) - \frac{1}{2}\right]^{b-1} + \phi\left(\frac{\varrho^{med}}{\sigma_{\varrho Pt}}\right)\left[\frac{1}{2} - \Phi\left(\frac{\varrho^{med}}{\sigma_{\varrho Pt}}\right)\right]^{a_c-1}\left[\Phi\left(\frac{\varrho^{med}+1}{\sigma_{\varrho Pt}}\right) - \frac{1}{2}\right]^{b}}{\left[\Phi\left(\frac{\varrho^{med}+1}{\sigma_{\varrho Pt}}\right) - \Phi\left(\frac{\varrho^{med}}{\sigma_{\varrho Pt}}\right)\right]^{a_c+b}} \tag{32}$$

for $\varrho^{med} \in (-1, 0)$, and where $\sigma_{\varrho Pt} = \frac{a_c}{a_c+b}\sqrt{\sigma_\zeta^2 + \Omega_{ct}^2 \frac{1}{\sigma_\xi^2} - 2\Omega_{ct}\rho\frac{\sigma_\zeta}{\sigma_\xi}}$.

[44]Notice that $\sigma_{\varrho Pt}^2 \to \sigma_{\varrho N}^2$ as $\Omega_t \to 0$, which implies that $\varrho_P^{med} \to_d \varrho_N^{med}$.

Replacing $\sigma_{\varrho N} = \frac{a_c}{a_c+b}\sigma_\zeta$ for $\sigma_{\varrho Pt}$ everywhere in (32), the density of unobserved heterogeneity in preferred law enforcement during periods under no Prohibition is obtained:$f_{\varrho_N^{med}}(\varrho_N^{med}; a_c, b_c, \sigma_{\varrho N})$.

□

Appendix 3: Derivation of the Conditional Likelihood

The joint density function of $(z_{ct}, \mu_{ct}, \varepsilon_{ct})$ is given by

$$f_{z\mu\varepsilon}(z_{ct}, \mu_{ct}, \varepsilon_{ct}; a_c, b, \overline{z}_{ct}, k, \lambda, \sigma_q^2, \sigma_z^2) = \frac{1}{\sqrt{2\pi}\sigma_z} exp\left(-\frac{1}{2\sigma_z^2}(z_{ct}-\overline{z}_{ct})^2\right) \frac{\mu_{ct}^{a_c-1}(1-\mu_{ct})^{b-1}}{\int x^{a_c-1}(1-x)^{b-1}dx} \frac{1}{\sqrt{2\pi}\sigma_q} exp\left(-\frac{\varepsilon_{ct}^2}{2\sigma_q^2}\right)$$

From (19), (20), and (21), $(z_{ct}, \mu_{ct}, \varepsilon_{ct})$ can be expressed as a function of the observables (p_{ct}, d_{ct}, q_{ct}):
From (21),

$$z_{ct} \equiv g_z(p_{ct}, d_{ct}, q_{ct}; \varrho_N^{med}, \varrho_P^{med}) = \frac{a_c+b}{a_c} \frac{1}{\alpha_{ct} k(\tau_{ct})} exp((\alpha_{ct}-1)p_{ct}+1) - \frac{a_c+b}{a_c}\left[P_{ct}\varrho_P^{med} + (1-P_{ct})\varrho_N^{med}\right] + P_{ct}\Omega_{ct}\overline{\theta}_{ct}^C - 1$$

From (20),

$$\mu_{ct} \equiv g_\mu(p_{ct}, d_{ct}, q_{ct}) = \frac{d_{ct}}{k(\tau_{ct})}(\chi + exp(\alpha_{ct} p_{ct}))$$

Finally from (19), and replacing for μ_{ct} from above,

$$\varepsilon_{ct} \equiv g_\varepsilon(p_{ct}, d_{ct}, q_{ct}) = q_{ct} - \Theta_S - d_{ct}(\chi + exp(\alpha_{ct} p_{ct}))\left\{exp(-\alpha_{ct} p_{ct}) + P_{ct}\theta_c\left[1 - exp(-\alpha_{ct} p_{ct})\right]\right\}$$

Now, iff $g(p_{ct}, d_{ct}, q_{ct}) = (g_z, g_\mu, g_\varepsilon)$ is a one-to-one mapping from (p_{ct}, d_{ct}, q_{ct}) to $(z_{ct}, \mu_{ct}, \varepsilon_{ct})$, the density function for (p_{ct}, d_{ct}, q_{ct}) will be given by

$$f_{pdq}(p_{ct}, d_{ct}, q_{ct}) = f_{z\mu\varepsilon}(g_z(p_{ct}, d_{ct}, q_{ct}; \varrho_N^{med}, \varrho_P^{med}), g_\mu(p_{ct}, d_{ct}, q_{ct}), g_\varepsilon(p_{ct}, d_{ct}, q_{ct}); a_c, b, \overline{z}_{ct}, k, \lambda, \sigma_q^2, \sigma_z^2)\,|J_{ct}|$$

where $|J_{ct}|$ is the absolute value of the determinant of the jacobian of g:

$$|J_{ct}| = \begin{vmatrix} \frac{\partial g_z}{\partial p_{ct}} & \frac{\partial g_z}{\partial d_{ct}} & \frac{\partial g_z}{\partial q_{ct}} \\ \frac{\partial g_\mu}{\partial p_{ct}} & \frac{\partial g_\mu}{\partial d_{ct}} & \frac{\partial g_\mu}{\partial q_{ct}} \\ \frac{\partial g_\varepsilon}{\partial p_{ct}} & \frac{\partial g_\varepsilon}{\partial d_{ct}} & \frac{\partial g_\varepsilon}{\partial q_{ct}} \end{vmatrix}$$

Given the structure of the model, conveniently $\frac{\partial g_z}{\partial d} = \frac{\partial g_z}{\partial q} = \frac{\partial g_\mu}{\partial q} = 0$, and $\frac{\partial g_\varepsilon}{\partial q} = 1$. To show that $g(p_{ct}, d_{ct}, q_{ct})$ is a one-to-one mapping, it is sufficient that $\frac{\partial g_z}{\partial p}$, $\frac{\partial g_\mu}{\partial d}$, $\frac{\partial g_\mu}{\partial p}$, $\frac{\partial g_\varepsilon}{\partial d}$, and $\frac{\partial g_\varepsilon}{\partial p}$ do not change sign. Solving for these derivatives,

$$\frac{\partial g_z}{\partial p} = \frac{a_c + b}{a_c} \frac{\alpha_{ct} - 1}{\alpha_{ct} k(\tau_{ct})} exp((\alpha_{ct} - 1)p_{ct} + 1)$$

which is always positive.

$$\frac{\partial g_\mu}{\partial d} = \frac{\chi + exp(\alpha_{ct} p_{ct})}{k(\tau_{ct})}$$

which is always positive.

$$\frac{\partial g_\mu}{\partial p} = \frac{d_{ct}}{k(\tau_{ct})} \alpha_{ct} exp(\alpha_{ct} p_{ct})$$

which is always positive.

$$\frac{\partial g_\varepsilon}{\partial d} = -(\chi + exp(\alpha_{ct} p_{ct})) \left\{ exp(-\alpha_{ct} p_{ct}) + P_{ct} \theta_c \left[1 - exp(-\alpha_{ct} p_{ct})\right] \right\}$$

which is always negative. Finally,

$$\frac{\partial g_\varepsilon}{\partial p} = -d_{ct} \alpha_{ct} \left[P_{ct} \theta_c \left[exp(\alpha_{ct} p_{ct}) + \chi exp(-\alpha_{ct} p_{ct}) \right] - \chi exp(-\alpha_{ct} p_{ct}) \right]$$

Notice that under no Prohibition, $\frac{\partial g_\varepsilon}{\partial p} > 0$ for any value of p_{ct}. Under Prohibition, a sufficient condition for $\frac{\partial g_\varepsilon}{\partial p} < 0$ (so that total crime is increasing in law enforcement) is that $\theta_c > \frac{\chi}{\chi + e_{ct}^{2\alpha} p_{ct}}$. In this case, $g(p_{ct}, d_{ct}, q_{ct})$ is one-to-one, and $|J|$ reduces to $|J| = \frac{\partial g_z}{\partial p} \frac{\partial g_\mu}{\partial d} \frac{\partial g_\varepsilon}{\partial q}$.

$$|J| = \frac{\partial g_z}{\partial p} \frac{\partial g_\mu}{\partial d}$$

Appendix 4: Additional Reduced Form Results

In this Appendix I discuss some additional reduced form results. First I present the results of the models whose coefficients are depicted in figures 4 and 5. Then I document the evolution of crime after Prohibition repeal, and look at neighboring alcohol markets and at the effect of pre-Prohibition Dry legislation and Women's suffrage, as alternative explanations for the main patterns described in section 4. Finally discuss the selection problem in the public opinion models.

Time-Varying Effects of Prohibition

Table A4-1 presents the main regression results from estimating equation (1). The tables present estimates of the specification including year effects for each outcome variable. As a benchmark for comparison, the first two columns for each outcome only include an indicator variable for years under Constitutional Prohibition (1920-1933). The next two columns then include the D_τ's instead of the Constitutional Prohibition indicator as a way of disaggregating the time-varying effects of Prohibition and allowing for pre-nationwide Prohibition effects[45]

Table A4-1: Long and Short-Run effects of Prohibition

Short and long-run effects of Prohibition

Dependent variable	Homicide Rate per 100,000		Drunkenness Arrest Rate per 1,000		Police Expenditure Share		Per Capita Police Expenditure	
	B	C	B	C	B	C	B	C
	(3)	(4)	(7)	(8)	(11)	(12)	(15)	(16)
1st Year under Prohibition	-2.290	-2.104	-7.765	-5.593	-0.0017	-0.0017	-0.032	-0.014
	(1.417)	(1.389)	(2.146)	(1.287)	(0.003)	(0.001)	(0.060)	(0.027)
2nd Year under Prohibition	-1.649	-1.252	-10.838	-9.737	-0.0025	-0.0017	0.014	0.003
	(1.840)	(1.848)	(2.538)	(1.432)	(0.003)	(0.002)	(0.083)	(0.037)
3rd Year under Prohibition	-1.369	-1.482	-10.947	-9.467	-0.0046	-0.0024	0.027	0.006
	(1.958)	(1.876)	(2.995)	(1.716)	(0.004)	(0.002)	(0.096)	(0.046)
4th Year under Prohibition	-3.123	-2.647	-11.721	-8.892	-0.0007	-0.0007	0.094	0.043
	(3.071)	(2.978)	(3.609)	(2.008)	(0.005)	(0.002)	(0.101)	(0.052)
5th Year under Prohibition	-0.072	0.281	-9.095	-7.859	0.0025	0.0002	0.149	0.075
	(2.439)	(2.459)	(3.844)	(2.310)	(0.005)	(0.003)	(0.101)	(0.055)
6th Year under Prohibition	0.330	0.738	-8.527	-7.422	0.0043	0.0024	0.213	0.122
	(2.818)	(2.785)	(4.216)	(2.623)	(0.005)	(0.003)	(0.129)	(0.058)
7th Year under Prohibition	0.117	0.731	-9.241	-7.777	0.0064	0.0046	0.243	0.155
	(2.820)	(2.746)	(4.676)	(2.601)	(0.006)	(0.003)	(0.163)	(0.062)
8th Year under Prohibition	2.381	2.937	-9.238	-8.326	0.0090	0.0058	0.351	0.181
	(2.331)	(2.395)	(4.676)	(2.720)	(0.005)	(0.003)	(0.177)	(0.074)
9th Year under Prohibition	2.320	3.163	-9.432	-8.083	0.0125	0.0070	0.422	0.213
	(2.628)	(2.709)	(4.996)	(2.801)	(0.006)	(0.003)	(0.196)	(0.081)
10th Year under Prohibition	1.528	2.397	-9.478	-7.908	0.0122	0.0072	0.412	0.262
	(2.417)	(2.553)	(5.574)	(2.963)	(0.006)	(0.003)	(0.189)	(0.063)
11th Year under Prohibition	0.841	1.853	-8.045	-7.698	0.0130	0.0074	0.456	0.269
	(2.323)	(2.505)	(5.644)	(2.995)	(0.005)	(0.002)	(0.181)	(0.059)
12th Year under Prohibition	-0.488	0.904	-5.891	-7.889	0.0164	0.0091	0.438	0.275
	(1.800)	(2.005)	(6.168)	(3.349)	(0.005)	(0.002)	(0.177)	(0.060)
13th Year under Prohibition	1.454	2.663	-4.241	-6.195	0.0158	0.0075	0.436	0.309
	(1.700)	(1.812)	(6.495)	(3.980)	(0.004)	(0.002)	(0.151)	(0.071)
14th Year under Prohibition	0.834	2.597	-1.574	-2.151	0.0123	0.0053	0.319	0.150
	(1.377)	(1.679)	(7.353)	(3.880)	(0.004)	(0.002)	(0.147)	(0.065)
15th Year under Prohibition	-2.906	-0.381	-6.240	-2.561	0.0083	0.0034	0.253	0.124
	(1.443)	(1.789)	(5.800)	(4.904)	(0.003)	(0.002)	(0.113)	(0.057)
16th Year under Prohibition	-3.146	-2.034	-2.467	-6.930	0.0064	0.0037	-0.016	-0.053
	(1.675)	(1.730)	(6.333)	(2.853)	(0.003)	(0.002)	(0.120)	(0.067)
17th Year under Prohibition	-4.601	-2.282	3.591	-10.118	0.0046	0.0025	-0.187	-0.137
	(1.605)	(1.655)	(6.486)	(3.336)	(0.004)	(0.003)	(0.100)	(0.088)
18th Year under Prohibition	-4.409	-3.033	1.493	-12.096	0.0008	-0.0004	-0.237	0.016
	(1.691)	(2.128)	(6.955)	(2.976)	(0.004)	(0.004)	(0.117)	(0.084)
19th Year under Prohibition	-7.033	-8.162	2.012	-8.127	0.0011	0.0039	-0.197	0.110
	(2.290)	(2.682)	(7.097)	(3.293)	(0.009)	(0.004)	(0.223)	(0.097)
Time-varying Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
City Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R squared	0.311	0.299	0.320	0.266	0.313	0.253	0.664	0.565
No. of Cities	66	90	66	237	66	239	66	239
No. of Observations	1716	1921	1254	4083	1650	4718	1650	4718

Notes: Constant included in all regressions is not reported. Standard errors are robust and clustered at the city level.
Time varying controls include log population, a Border indicator and a State-capital indicator.
Sample B is the balanced sample used for Structural estimation. Sample C includes all cities for which at least 8 years of data are available.

Effects of Heterogeneity in Moral Views during Prohibition

Table A4-2 now presents the estimated models from equation (3). Results are presented following the same structure as those in table A4-1, reporting the interaction terms only.

[45]For the police share and per capita police expenditure regressions, I also ran regressions including city-specific trends, which do not show any significative differences to the ones presented in table A4-1.

Table A4-2: Moral Heterogeneity: Long-and Short Run Differences

	Differences in Moral Views							
Dependent variable	Homicide Rate per 100,000		Drunkenness Arrest Rate per 1,000		Police Expenditure Share		Per Capita Police Expenditure	
	B	C	B	C	B	C	B	C
	(3)	(4)	(7)	(8)	(11)	(12)	(15)	(16)
1st Year under Prohibition x Wetness	-13.584	-13.161	-10.986	7.387	-0.097	-0.036	-3.304	-1.178
	(14.735)	(14.296)	(21.368)	(8.716)	(0.028)	(0.012)	(0.855)	(0.315)
2nd Year under Prohibition x Wetness	-2.103	1.406	-13.652	12.240	-0.100	-0.032	-3.478	-0.943
	(11.105)	(9.480)	(22.109)	(11.701)	(0.032)	(0.014)	(0.943)	(0.318)
3rd Year under Prohibition x Wetness	-5.829	-3.868	-5.762	19.051	-0.103	-0.016	-3.400	-0.635
	(9.234)	(8.489)	(23.055)	(12.403)	(0.037)	(0.016)	(1.068)	(0.340)
4th Year under Prohibition x Wetness	9.867	8.950	3.567	23.345	-0.106	-0.011	-3.237	-0.520
	(8.692)	(8.489)	(20.371)	(11.902)	(0.035)	(0.015)	(1.027)	(0.351)
5th Year under Prohibition x Wetness	8.545	11.051	-0.761	25.510	-0.107	-0.004	-3.158	-0.388
	(7.449)	(7.657)	(20.323)	(11.646)	(0.034)	(0.016)	(1.023)	(0.392)
6th Year under Prohibition x Wetness	17.975	21.918	6.486	14.687	-0.097	-0.006	-2.775	-0.349
	(9.142)	(9.429)	(20.533)	(12.512)	(0.028)	(0.015)	(0.923)	(0.430)
7th Year under Prohibition x Wetness	28.179	26.301	0.257	9.445	-0.093	-0.003	-2.144	-0.072
	(11.930)	(7.865)	(21.481)	(11.561)	(0.030)	(0.015)	(0.961)	(0.458)
8th Year under Prohibition x Wetness	20.584	23.788	-7.061	3.815	-0.071	0.004	-2.143	-0.118
	(7.737)	(7.566)	(22.304)	(11.489)	(0.026)	(0.016)	(0.827)	(0.519)
9th Year under Prohibition x Wetness	32.922	32.455	-11.510	0.992	-0.055	0.003	-1.240	-0.385
	(11.015)	(9.942)	(22.904)	(11.363)	(0.024)	(0.015)	(0.879)	(0.642)
10th Year under Prohibition x Wetness	22.065	22.236	-3.044	1.578	-0.059	0.005	-1.520	-0.008
	(9.146)	(8.214)	(22.696)	(11.146)	(0.022)	(0.016)	(0.795)	(0.567)
11th Year under Prohibition x Wetness	29.034	29.361	-4.601	2.361	-0.036	0.014	-0.843	0.201
	(10.497)	(9.270)	(23.652)	(13.084)	(0.020)	(0.016)	(0.780)	(0.580)
12th Year under Prohibition x Wetness	27.208	28.249	28.040	-1.229	0.036	0.024	1.273	0.908
	(7.488)	(6.853)	(28.151)	(22.000)	(0.020)	(0.014)	(0.759)	(0.511)
13th Year under Prohibition x Wetness	17.276	15.742	18.806	-21.000	0.014	0.023	1.741	1.625
	(7.821)	(6.388)	(39.119)	(28.344)	(0.018)	(0.014)	(0.929)	(0.519)
14th Year under Prohibition x Wetness	19.820	15.645	28.856	6.424	0.009	0.033	2.353	1.590
	(6.940)	(6.432)	(83.417)	(43.458)	(0.018)	(0.016)	(0.985)	(0.629)
15th Year under Prohibition x Wetness	-1.986	0.901	-110.498	-214.430	0.003	0.020	2.622	1.753
	(9.616)	(7.330)	(74.038)	(47.447)	(0.023)	(0.017)	(1.022)	(0.728)
16th Year under Prohibition x Wetness	17.982	6.842	-4.001	-46.925	0.008	0.034	1.382	-0.374
	(15.128)	(13.627)	(66.963)	(46.322)	(0.039)	(0.024)	(1.597)	(0.859)
17th Year under Prohibition x Wetness	-4.708	4.993		172.815	0.027	0.005	0.217	-0.734
	(8.743)	(7.385)		(135.191)	(0.019)	(0.032)	(0.401)	(1.046)
18th Year under Prohibition x Wetness	8.580	11.885		-22.205	0.011	0.116	0.295	0.232
	(7.886)	(7.348)		(72.675)	(0.025)	(0.093)	(0.538)	(1.126)
19th Year under Prohibition x Wetness	7.646	14.108		-64.025	-0.014	-0.044	0.296	-1.724
	(9.716)	(10.669)		(50.044)	(0.040)	(0.068)	(1.003)	(1.692)
Time-varying Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
City Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R squared	0.337	0.326	0.323	0.273	0.358	0.262	0.690	0.574
No. of Cities	66	90	66	237	66	239	66	239
No. of Observations	1716	1921	1254	4083	1650	4718	1650	4718

Notes: Constant included in all regressions is not reported. Standard errors are robust and clustered at the city level.
Time varying controls include log population, a Border indicator and a State-capital indicator.
Sample B is the balanced sample used for Structural estimation. Sample C includes all cities for which at least 8 years of data are available.

Prohibition Repeal

The repeal of the 18th Amendment itself also allows for the exploration of differential trends in criminality between cities with varying moral profiles. Here I exploit the repeal of nationwide Prohibition in December 1933 with the ratification of the 21st Amendment, to provide some additional evidence of the response of crime to Prohibition, and its stronger effects in communities with larger alcohol markets. I take advantage of the availability of more detailed crime data for the 1930-1936 period, taken from the Uniform Crime Reports (UCR) complied by the FBI starting in 1930. The UCR reports for a large number of cities, the total number of offences known to the authorities (which include any of the following: murder, rape, robbery, assault, burglary, larceny, and auto theft), and an independent measure of reported murders. Thus, I compare crime outcomes in the 1930-1933 period with the 1934-1936 period, allowing for differential behavior after repeal, as cities vary in their moral preference distribution, as proxied by μ. Indeed, simple summary statistics show that offences and murders were both lower in the post-18th Amendment years [46].

Thus, I look exclusively at the period 1930-1936, and run regressions for UCR offences and murders, and for the homicide rate.

$$y_{ct} = \alpha_c + \beta_t + \delta CP_t + \phi CP_t \overline{\mu}_c + \boldsymbol{\gamma}' \boldsymbol{X}_{ct} + \varepsilon_{ct} \tag{33}$$

where CP_t is an indicator variable for Constitutional Prohibition. Regression results are reported in table A4-3. Columns $(1)-(4)$ look at the homicide rate. The coefficient on the interaction is always large, highly significant, and robust to the introduction of state-cross-year effects, suggesting that the fall in crime was larger in wetter cities. Take for example column (2). The estimates imply that for the city with mean "wetness" of 0.49, repeal was associated with a fall in the annual homicide rate of $4.6 = (0.49 \times 23.28) - 6.76$. Even in the driest city, with $\mu = 0.3$, the estimated fall in the homicide rate is $0.21 = (0.3 \times 23.28) - 6.76$. Columns $(5)-(8)$ then present analogous results for the UCR number of murders per $100,000$. The pattern is very similar to the one for the homicide rate, although standard errors increase somewhat, and the magnitude of the effect is smaller for the larger sample of cities covered. Nonetheless, for the sample for which homicide rates are available, results are very similar. The large standard errors for the sample in Columns (6) and (8) is due to the larger number of smaller cities included, in which reported murders were very small or close to zero, and present very little variation. Finally, Columns $(9)-(12)$ present results for offences per $1,000$. Interestingly, a pattern very similar to the one for homicides and arrests emerges, but this time, the effect is statistically significant especially in the larger sample including cities of smaller sizes. From Column (12), for example, it follows that repeal in the city with average "wetness" implied a fall in total offences of $3.85 = (0.49 \times 6.53) + 0.669$ per $1,000$ population, which is 43% of this variable's standard deviation of 8.65. As the results suggest, while the reduction in criminality in larger cities was associated especially with a lower homicide rate, looking at a larger sample including smaller cities, repeal was associated with lower levels of other types of crime.

Table A4-3: Crime Fall after Repeal

[46] Average murders per $100,000$ are 8.57 (s.e. = 10.3) in the 1930-1933 period, and 6.53 (s.e. = 8.8) in 1934-1936, with a t-statistic for the difference in means of 4.62. For offences per $1,000$, the 1930-1933 mean is 16.26 (s.e. = 8.6), while the 1934-1936 mean is 15.6 (s.e. = 8.6), with a t-statistic of 1.64, significant at the 5% level.

	Repeal of the 18th Amendment											
Dependent variable	Homicide Rate per 100,000				Murders per 100,000				All Offences per 100,000			
	B	C	B	C	B	C	B	C	B	C	B	C
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Constitutional Prohibition (1930-1933)	-5.91	-6.77	-6.01	-4.35	1.27	-0.44	9.14	3.23	-0.46	-3.95	-4.12	0.67
	(4.18)	(3.74)	(5.39)	(3.53)	(5.79)	(2.00)	(10.49)	(9.06)	(4.88)	(1.84)	(7.45)	(2.53)
Constitutional Prohibition x Wetness	20.12	23.28	22.13	18.47	4.04	3.88	15.55	3.37	3.04	5.32	12.22	6.53
	(7.42)	(6.65)	(11.93)	(7.80)	(10.93)	(3.64)	(23.20)	(5.48)	(8.53)	(3.21)	(16.48)	(3.88)
Time-varying Controls	Yes	Yes	No	No	Yes	Yes	No	No	Yes	Yes	No	No
Year Effects	Yes	Yes	No	No	Yes	Yes	No	No	Yes	Yes	No	No
City Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
State x Year Effects	No	No	Yes	Yes	No	No	Yes	Yes	No	No	Yes	Yes
R squared	0.450	0.353	0.713	0.622	0.349	0.085	0.609	0.249	0.121	0.067	0.476	0.2307
No. of Cities	66	93	66	93	66	324	66	324	66	324	66	324
No. of Observations	462	651	462	651	417	1938	417	1938	414	1943	414	1943

Notes: Constant included in all regressions is not reported. Standard errors are robust and clustered at the city level.
Time varying controls include log population, a Border indicator and a State-capital indicator. Each specification is estimated for three samples:
Sample A includes all city-x-year observations for which data is available. Sample B is the balanced sample used for Structural estimation.
Sample C includes all cities for which at least 10 years of data are available.

Neighboring Markets

If individuals' preferences are affected by the legal standard in place, say because they derive utility from abiding to the law, or, on the other hand, if individuals' utility from taking an action increases when it is proscribed (a "forbidden fruit effect"), observed changes in the drunkenness arrest rate could be driven by these shocks in preferences. A way to isolate any taste shocks introduced by Prohibition is to look at the response of the alcohol market in a city which is already under Prohibition, when neighboring states' prohibitionist status changes. If drinkers in a city under Prohibition have access to neighboring markets, which is very consistent with the concern of Prohibitionists of the time, and which motivated the passage of the Webb-Kenyon act, then the closure of neighboring markets should reduce the availability of liquor in the city, without having an effect on preferences[47]. Thus, I collected information on the lengths of all state boundaries[48], and computed for each state, the share of state border in states under Prohibition at each point in time[49]:

$$SBP_{ct} = \frac{\sum_{j \in N_c} P_{jt} \times BorderLength_{cj}}{\sum_{j \in N_c} BorderLength_{cj}}$$

where P_{jt} is an indicator variable for state j being under Prohibition at time t. N_c is the set of states neighboring city c's state (e.g. $N_{SanFrancisco} = \{Oregon, Nevada, Arizona, Mexico\}$), and $BorderLength_{cj}$ is the length in miles of the state boundary between city c's state and state j.

[47] The importance of cross-state-boundaries alcohol trade after Prohibition was enacted in some states but not in neighboring ones is probably best exemplified by Daniel Okrent's discussion of the huge traffic lanes along interstate 25, connecting Toledo, OH with Detroit, MI, after Michigan was covered by state-wide Prohibition in 1918. The highway was nicknamed "Avenue de Booze". (See Okrent (2010, p. 107))

Isaac also highlights the importance of cross-state smuggling of alcohol after Tennessee started enforcing its Prohibition legislation: "The State bone-dry law, even when supplemented by the Reed amendment, or "national bone-dry law", which made it a federal crime to transport intoxicants into a dry state, did not actually stop the flow of liquor into Tennessee. During 1917 and 1918, bootleggers where adequately supplied with whiskey brought from Kentucky to Nashville and Memphis by train, automobile, farm wagon, and river boat." Isaac (1965, p. 254)

[48] The information on state boundary lengths was taken from Holmes (1996). There are a total of 109 boundaries between U.S. states, and 16 international boundaries.

[49] I include any international borders in the denominator, which amounts to considering Mexico and Canada as never being under Prohibition.

For the pre-Constitutional Prohibition period (1911-1919), when there is variation across states in Prohibition status, I estimate models of the form

$$d_{ct} = \alpha_c + \beta_t + \delta P_{ct} + \eta SBP_{ct} + \phi P_{ct} SBP_{ct} + \boldsymbol{\gamma}' \boldsymbol{X}_{ct} + \varepsilon_{ct} \tag{34}$$

Table A4-4 presents the estimates of equation (34), for different specifications, and for samples A, B, and C. First, the fraction of border under Prohibition should have an effect on the drunkenness arrest rate only when the city itself is under Prohibition; otherwise the city's neighbors' Prohibition status should be irrelevant, since a free alcohol market is available. Thus, columns $(1)-(3)$ in table A4-4 start presenting the estimates of a model where I include the share of border under Prohibition without an interaction with own Prohibition status. The share of state boundary under Prohibition is insignificant in the three specifications. Then columns $(4)-(6)$ introduce the interaction term, and columns $(7)-(9)$ additionally include time-varying controls (log of population, and time-varying state capital and South effects). The coefficient for the ϕ is negative and large in magnitude, and always highly significant, except for column (8) when looking at the smaller B sample. The coefficient for ϕ on column (6), for example, implies that a one pre-1920 standard deviation (0.29) increase in the fraction of state border under Prohibition implied a reduction in the drunkenness arrest rate of 1.93, which is 10% of the average pre-1920 drunkenness arrest rate in the sample. These estimates are very consistent with the idea that the sharp falls in drunkenness arrests observed were caused by a contraction in the alcohol supply available, and not due to preference shocks correlated with the introduction of Prohibition.

Table A4-4: Neighboring Prohibition

Effect of Neighboring Prohibition

Dependent variable	Drunkenness Arrests Rate per 1,000								
	A	B	C	A	B	C	A	B	C
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Prohibition Indicator	-6.123	-6.072	-6.415	-3.427	-5.464	-3.745	-3.511	-5.750	-3.932
	(1.376)	(1.927)	(1.359)	(1.864)	(2.182)	(1.859)	(2.165)	(2.652)	(2.149)
Share of Border under Prohibition	-2.128	-6.257	-2.102	1.455	-5.504	1.425	0.351	-6.144	0.300
	(2.138)	(3.095)	(2.139)	(2.562)	(3.338)	(2.562)	(2.440)	(3.167)	(2.441)
Prohibition X Share of Border under Prohibition				-6.778	-1.348	-6.681	-5.112	-0.046	-4.976
				(2.650)	(3.018)	(2.656)	(2.819)	(3.685)	(2.830)
Time-varying Controls	No	No	No	No	No	No	Yes	Yes	Yes
Year Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
City Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R squared	0.226	0.319	0.228	0.234	0.320	0.236	0.248	0.353	0.251
No. of Cities	245	66	236	245	66	236	245	66	236
No. of Observations	1876	594	1861	1876	594	1861	1876	594	1861

Notes: Constant included in all regressions is not reported. Standard errors are robust and clustered at the city level. Time varying controls include log population, a Border indicator and a State-capital indicator. Each specification is estimated for three samples: Sample A includes all city-x-year observations for which data is available. Sample B is the balanced sample used for Structural estimation. Sample C includes all cities for which at least 10 years of data are available. Coefficients for years under Prohibition and for the interactions between "wetness" and years under Prohibition not reported.

Dry Legislation

Can differences in pre-Prohibition alcohol-related legislation explain the trends in crime, arrests for drunkenness, and police enforcement? Prior to the adoption of state-level and nationwide Prohibition, different states had different types and numbers of dry laws. In fact, regulations over the alcohol market were in place almost everywhere. These included restrictions on selling hours, on the kinds of alcoholic beverages permitted, on the types of selling establishments allowed, and on taxation. There are two channels through which pre-Prohibition alcohol legislation might affect the evolution of outcomes during Prohibition. First, given that early on during Prohibition collective law enforcement decisions were likely to be closely related to initial "prior" beliefs about the policy's effects, variation in the short-run effects of Prohibition might be partly explained by variation in pre-Prohibition dry legislation. The direction of an effect is not obvious *a priori*. On the one hand, if these laws were being successful in shrinking the alcohol market and were not affecting crime, people's priors about the introduction of federal-level Prohibition could be very optimistic; on the other hand, if the introduction of these laws was correlated with more crime, individuals might have used this information to form negative priors about nationwide Prohibition. Second, differences in dry laws could have created different initial conditions for the alcohol market at the time of Prohibition adoption. For example, heavily regulated markets might have already developed a parallel black market which could have eased the expansion of the illegal liquor trade during Prohibition.

To take a look at this question I reviewed the available information on state-level dry legislation in the pre-18th Amendment period and constructed a variable counting the number of regulations on the alcohol market at each point in time for each state. Interestingly, although the relationship between average "wetness" of a state, as measured by μ, and the number of dry laws in place is not very strong, it is actually positive. This is likely to be the result of the equilibrium political strategies used by dry lobbies during the 1900s and 1910s. Because relatively "wet" regions were unlikely to pass Prohibition laws, the lobbies focused their efforts on passing regulatory legislation instead, which was politically feasible[50]. States like Michigan or Minnessota (both heavily "wet"), passed, especially during the 1910s, significant amounts of regulatory legislation related to alcohol. In the other extreme, radically "dry" states such as Utah and Oklahoma did not need to pass this kind of legislation because they were already under Prohibition in the first place.

Pre-Prohibiton legislation is, of course, endogenous to outcomes over that period. Given that I want to explore the effects of pre-Prohibition dry legislation on outcomes during Prohibition, which might have an effect through initial beliefs (and hence, initial law enforcement choices during Prohibition), or in how they shaped the local alcohol markets (and hence, in the subsequent response of alcohol supply during Prohibition), below I briefly investigate the effect of pre-Prohibition legislation on Prohibition outcomes, conditional on local preferences, by estimating models only for Prohibition years, in which I allow for a differential effect of the number of pre-Prohibition dry laws over time under Prohibition, controlling by a time-varying effect of baseline "wetness":

$$y_{ct} = \alpha_c + \beta_t + \sum_{\tau=1}^{k} \delta_\tau D_{c\tau} + \sum_{\tau=1}^{k} \eta_\tau D_{c\tau} L_c + \boldsymbol{\gamma}' \boldsymbol{X}_{ct} + \varepsilon_{ct} \tag{35}$$

where L_c is the number of dry laws in place right before the city is under Prohibition, and $\boldsymbol{X}_{ct}$ includes interactions of $\overline{\mu}_{c0}$ with year indicators. Because these models only look at years under Prohibition, I omit the indicator for $\tau = 1$, so the interpretation of the "years under Prohibition" indicator variables is different; coefficients must now be interpreted as relative to having experienced Prohibition for one year. The η_τ's should capture any time-varying differential effects of an extra piece of dry pre-Prohibition legislation on Prohibition outcomes. Flexibly controlling for the moral

[50]The data on dry legislation was mostly taken from Cherrington (1920)and League (1932). Both sources have a detailed and comprehensive compilation of dry legislation during these decades.

profile of the city as proxied by μ is important given that pre-Prohibition dry legislation is likely to be correlated with preferences in the city. To save space, in table A4-5 I only present results for the coefficient estimates for the η_τ's of the benchmark fixed effects specifications. Regression results do not show any significant relationship between the amount of pre-Prohibition dry legislation and the homicide rate or the arrest rate at any time during Prohibition. There also appears to be no relation between these laws and the behavior of per capita expenditure in policing during Prohibition years. For the expenditure share, on the other hand, the interaction terms are small in magnitude but significant, suggesting up to a 1% higher police share per pre-Prohibition piece of legislation around the 10th year under Prohibition, relative to the first one (See column (6)). This result is not robust to the introduction of city-specific trends, though. Overall there seems to be no evidence that dry laws prior to Prohibition had any economically important effects on the evolution of outcomes during Prohibition years.

Table A4-5: Effects of pre-Prohibition Dry Legislation

Dry Laws								
Dependent variable	Homicide Rate per 100,000		Drunkenness Arrests Rate per 1,000		Police Expenditure Share		Per Capita Police Expenditure	
	B	C	B	C	B	C	B	C
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
2nd Year under Prohibition x Pre Prohibition Dry Laws	0.098	0.166	-0.334	-0.189	-0.0005	0.0001	-0.004	-0.004
	(0.208)	(0.202)	(0.398)	(0.291)	(0.0003)	(0.0002)	(0.006)	(0.004)
3rd Year under Prohibition x Pre Prohibition Dry Laws	0.247	0.218	-0.237	0.120	-0.0006	0.0001	-0.008	-0.005
	(0.200)	(0.195)	(0.409)	(0.285)	(0.0005)	(0.0003)	(0.009)	(0.006)
4th Year under Prohibition x Pre Prohibition Dry Laws	0.452	0.469	-0.316	0.163	-0.0004	0.0002	-0.013	-0.007
	(0.229)	(0.229)	(0.407)	(0.282)	(0.0006)	(0.0003)	(0.011)	(0.007)
5th Year under Prohibition x Pre Prohibition Dry Laws	0.183	0.174	-0.427	0.122	0.0000	0.0005	-0.018	-0.010
	(0.289)	(0.286)	(0.435)	(0.280)	(0.0006)	(0.0003)	(0.013)	(0.008)
6th Year under Prohibition x Pre Prohibition Dry Laws	0.319	0.289	-0.503	-0.139	0.0005	0.0008	-0.005	-0.005
	(0.262)	(0.262)	(0.427)	(0.290)	(0.0007)	(0.0003)	(0.019)	(0.009)
7th Year under Prohibition x Pre Prohibition Dry Laws	0.334	0.315	-0.431	-0.129	0.0004	0.0009	-0.008	-0.003
	(0.312)	(0.304)	(0.451)	(0.296)	(0.0007)	(0.0003)	(0.021)	(0.010)
8th Year under Prohibition x Pre Prohibition Dry Laws	0.375	0.363	-0.486	-0.010	0.0009	0.0011	-0.003	-0.001
	(0.307)	(0.299)	(0.413)	(0.290)	(0.0007)	(0.0003)	(0.022)	(0.010)
9th Year under Prohibition x Pre Prohibition Dry Laws	0.353	0.369	-0.672	-0.054	0.0012	0.0012	0.005	0.001
	(0.327)	(0.318)	(0.428)	(0.306)	(0.0008)	(0.0003)	(0.025)	(0.012)
10th Year under Prohibition x Pre Prohibition Dry Laws	0.298	0.321	-0.559	-0.056	0.0009	0.0011	-0.009	0.001
	(0.249)	(0.243)	(0.431)	(0.312)	(0.0007)	(0.0004)	(0.024)	(0.013)
11th Year under Prohibition x Pre Prohibition Dry Laws	0.094	0.125	-0.540	-0.188	0.0011	0.0011	-0.010	-0.013
	(0.251)	(0.245)	(0.418)	(0.294)	(0.0007)	(0.0004)	(0.024)	(0.019)
12th Year under Prohibition x Pre Prohibition Dry Laws	0.194	0.149	-0.394	-0.031	0.0008	0.0012	-0.022	-0.001
	(0.253)	(0.240)	(0.480)	(0.340)	(0.0008)	(0.0004)	(0.030)	(0.014)
13th Year under Prohibition x Pre Prohibition Dry Laws	0.133	0.094	-0.079	0.188	0.0010	0.0011	-0.034	-0.012
	(0.236)	(0.227)	(0.878)	(0.661)	(0.0009)	(0.0004)	(0.035)	(0.019)
14th Year under Prohibition x Pre Prohibition Dry Laws	0.149	0.196	-1.212	0.606	0.0012	0.0014	-0.025	0.005
	(0.263)	(0.250)	(1.081)	(1.197)	(0.0009)	(0.0004)	(0.037)	(0.018)
15th Year under Prohibition x Pre Prohibition Dry Laws	-0.019	0.157	-16.505	1.956	0.0015	0.0015	-0.024	0.004
	(0.242)	(0.245)	(3.514)	(2.523)	(0.0010)	(0.0005)	(0.038)	(0.021)
16th Year under Prohibition x Pre Prohibition Dry Laws	-0.143	0.055	-12.068	2.232	0.0021	0.0006	0.017	-0.002
	(0.440)	(0.367)	(1.764)	(1.429)	(0.0011)	(0.0009)	(0.038)	(0.025)
17th Year under Prohibition x Pre Prohibition Dry Laws	-0.913	-0.737	5.239	-1.343	0.0020	-0.0003	0.004	-0.053
	(1.304)	(1.035)	(1.039)	(5.556)	(0.0041)	(0.0022)	(0.078)	(0.046)
18th Year under Prohibition x Pre Prohibition Dry Laws	1.369	1.121	4.310	0.270	0.0008	-0.0025	-0.027	-0.056
	(1.404)	(1.492)	(1.077)	(4.264)	(0.0037)	(0.0030)	(0.072)	(0.053)
19th and more Years under Prohibition x Pre Prohibition Dry Laws	-4.095	-3.993	6.214	0.041	-0.0085	-0.0029	-0.465	-0.080
	(1.812)	(1.605)	(1.225)	(4.071)	(0.0053)	(0.0022)	(0.146)	(0.070)
Time-varying Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
City Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R squared	0.465	0.399	0.502	0.345	0.393	0.212	0.746	0.607
No. of Cities	66	90	66	237	66	239	66	239
No. of Observations	1066	1210	1066	2655	1066	3477	1066	3477

Notes: Constant included in all regressions is not reported. Standard errors are robust and clustered at the city level. Time varying controls include log population, a Border indicator and a State-capital indicator.
Sample B is the balanced sample used for Structural estimation. Sample C includes all cities for which at least 8 years of data are available.
Coefficients for years under Prohibition and for the interactions between "wetness" and years under Prohibition not reported.

Women's suffrage

Several historians have attributed some of the success of Prohibition in the United States to the significant role that the Women's Suffrage Movement played. It is undeniable that women played a prominent role in the conflict over alcohol consumption, and were of importance at least since the 1870s, when a group of Ohio women began the "Temperance Crusade" that spread throughout all of the Midwest. A group of women would visit the area's saloons one by one, and protest and pray for days until the owners decided to close. The long-term effects of the crusade are likely to have been minimal, but it was the first major women-specific social mobilization, and was the origin of the WCTU some years later. In the Twentieth century, both the Women's Suffrage Movement and the Temperance Movement were part of the Progressive-era reforms, and organizations such as the WCTU were involved in the political struggle around both issues. Although U.S.-wide women's suffrage (19th Amendment) was ratified into the Constitution in 1920, after the adoption of nationwide Prohibition (18th Amendment), authors such as Okrent (2010) argue that the Women's Suffrage Movement gave a major impulse to the Prohibition movement. The almost simultaneous ratification of the 18th and 19th Amendments makes it impossible to identify any specific effects that women's suffrage might have had during federal Prohibition years. Nonetheless, prior to the 19th Amendment several states had already extended the franchise to women [51]. As a way to explore the importance of women's enfranchisement on Prohibition-related outcomes, I exploit the variation in women's suffrage enfranchisement prior to 1920, when both the 18th and 19th Amendments were ratified, to see if Prohibition had differential effects in cities with and without women's suffrage. If the distribution of women's preferences over Prohibition enforcement was different than men's, cities allowing women's suffrage could be under a differential trend. Thus, for the 1910-1919 period, I run regressions of the form

$$y_{ct} = \alpha_c + \beta_t + \eta W_{ct} + \sum_{\tau=1}^{k} \delta_\tau D_\tau + \sum_{\tau=1}^{k} \phi_\tau D_\tau W_{ct} + \boldsymbol{\gamma}' \boldsymbol{X}_{ct} + \varepsilon_{ct} \tag{36}$$

In equation (36), W_{ct} is an indicator variable taking the value of 1 if city c has women's suffrage in year t. Table A4-6 presents results of the estimates of the ϕ_τ's from equation (36) for the different outcome variables, in the specifications including city fixed effects, time-varying controls, and year effects. The regressions include only up to ϕ_5, because before 1919 no city with women's suffrage in the sample had experienced more than 5 years under Prohibition. There is no evidence of a differential trend in the homicide rate in cities with women's suffrage. This is unsurprising given that the short-run effects of Prohibition on the homicide rate were very small. For the outcomes which did have large short-run changes after the introduction of Prohibition, if anything, Columns $(2)-(3)$ in table A4-6 show that the introduction of women's suffrage is correlated with more drunkenness arrests in the short run (after two to three years under Prohibiton), but the net effect is small and insignificant quickly thereafter. This result is also not robust to the restricted B sample (column (2)). When looking at police enforcement in Columns $(4)-(7)$, the results are also very inconclusive.

[51]Women's Suffrage prior to the 19th Amendment was adopted by the states as follows: Wyoming in 1869, Colorado in 1893, Utah and Idaho in 1896, Washington in 1910, California in 1911, Arizona, Kansas and Oregon in 1912, Montana and Nevada in 1914, New York in 1917, and Michigan, Oklahoma and South Dakota in 1918.

During years with women's suffrage, cities have slightly lower but insignificant policing, which is actually inconsistent with the idea that women's anti-Prohibitionism should translate to higher law enforcement and a smaller alcohol market after their enfranchisement. Overall, the available evidence does not suggest that alternative legislation, such as dry laws or women's suffrage, might have been driving the trends in law enforcement, crime and arrests presented in Section 4.1.

Table A4-6: Effect of Women's Suffrage

	Women's Suffrage						
Dependent variable	Homicide Rate per 100,000	Drunkenness Arrests Rate per 1,000		Police Expenditure Share		Per Capita Police Expenditure	
	B	B	C	B	C	B	C
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Women's Suffrage Indicator	0.1171	-1.080	-2.183	0.000	0.001	0.013	0.051
	(1.024)	(1.696)	(1.132)	(0.003)	(0.002)	(0.072)	(0.046)
1st Year under Prohibition x Women's Suffrage	-0.7345	1.410	1.311	-0.007	-0.003	-0.178	-0.085
	(1.502)	(2.812)	(1.623)	(0.003)	(0.002)	(0.092)	(0.040)
2nd Year under Prohibition x Women's Suffrage	-2.3699	-5.484	5.409	-0.017	-0.002	-0.124	0.002
	(4.357)	(3.447)	(2.645)	(0.006)	(0.004)	(0.110)	(0.054)
3rd Year under Prohibition x Women's Suffrage	-1.3185	-4.172	5.616	-0.021	-0.002	-0.151	-0.024
	(4.537)	(3.969)	(3.655)	(0.009)	(0.005)	(0.111)	(0.062)
4th Year under Prohibition x Women's Suffrage	5.5515	-0.017	4.276	-0.036	-0.003	-0.223	-0.051
	(6.904)	(6.215)	(3.767)	(0.010)	(0.005)	(0.095)	(0.051)
5th Year under Prohibition x Women's Suffrage			3.329				
			(4.841)				
Time-varying Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
City Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R squared	0.108	0.344	0.291	0.330	0.195	0.649	0.606
No. of Cities	66	66	236	66	217	66	217
No. of Observations	594	594	1861	528	1427	528	1427

Notes: Constant included in all regressions is not reported. Years Under Prohibition Indicators not reported.
Standard errors are robust and clustered at the city level. Time varying controls include log population, a Border indicator and a State-capital indicator.
Sample B is the balanced sample used for Structural estimation. Sample C includes all cities for which at least 10 years of data are available.

Selection in the Public Opinion data

A caveat in the elections data is that several states including Louisiana, New Jersey, New York, and Pennsylvania, did not hold any liquor-related referendum in a pre-Prohibition year. This induces a potential selection bias in the estimates of equation (4) because these states never held a referendum regarding liquor precisely due to the highly anti-Prohibitionist preferences of their citizens. As a robustness check I also estimate a selection model, by specifying a selection equation for holding a referendum (at the state level). More especifically, I assume that

$$r_{St} = \begin{cases} 1 & if\ t = 0\ and\ \boldsymbol{\eta'}\boldsymbol{Z}_{S0} + \upsilon_{S0} > 0 \\ 1 & if\ t = 1 \end{cases}$$

where r_{St} is an indicator variable for state S holding a liquor referendum, $\boldsymbol{Z}_{S0}$ includes the state's share of adherants to a wet religion and the share of native white individuals in 1910, and $\upsilon_{S0} \sim N(0,1)$, with $E[\varepsilon_{c0}|\upsilon_{S0}] = \rho\upsilon_{S0}$ and $E[\varepsilon_{c1}|\upsilon_{S0}] = 0$. This implies that

$$E[w_{ct}|\mu_{ct}, \mu_{c0}, \boldsymbol{X}_{ct}, r_{St} = 1] = \alpha_c + \beta t + \delta\mu_{ct} + \phi\mu_{c0}t + \boldsymbol{\gamma'}\boldsymbol{X}_{ct} + \kappa\lambda(\boldsymbol{\eta'}\boldsymbol{Z}_{S0})1_{\{t=0\}}$$

where $\lambda()$ is the inverse Mills ratio. Results are reported in columns (6), (12), and (18) of table 3.

Appendix 5: Data Sources

Most of the information available for the study of Prohibition is available at the city level, so I focused on constructing a yearly panel dataset of cities, covering the 1910s, '20s and early '30s. The data collected comes from a wide array of sources. The first source of information is the collection of original documents from the *National Commission on Law Observance and Enforcement*, most commonly known as the Wickersham Commission after the name of its Chair Commissioner, Attorney General George Wickersham. It was appointed in the Spring of 1929 by President Hoover, with the specific purpose of "studying exhaustively the entire problem of the enforcement of our laws and the improvement of our judicial system, including the special problem and abuses growing out of the Prohibition laws" (Wickersham-Commission (1928-1931b)). It was, of course, appointed as a response to the growing concerns about the effects Prohibition was having throughout the country, and the public discontent over the policy's effects. The Commission produced a series of reports on the different aspects of Prohibition, after directly collecting data and evidence across the country, and issued its main findings in 1931. Harvard's Law School Library curretly holds the collection of documents from the Commission, including the originals of much of the summarized data in the published reports, in addition to several other unpublished information. The detailed city-by-city "Prohibition Survey" reports, compiled directly by commissioners traveling to the cities and collecting information about the recent evolution of criminality, and the "Cost of Crime" state-level folders, providing detailed data on local law enforcement activity, contain the most valuable information from the Wickersham papers.

Law Enforcement Data

Other than the data mentioned in section 3, the Wickersham Commission papers also contain other data on total arrests, unfortunately, available only during the 1910s and in 1929. Data on a set of other Prohibition enforcement outcomes is available only at the state level from the U.S. Bureau of Prohibition for the years 1923-1932, such as the number of still and liquor seizures, arrests of alcohol producers, and casualties caused by Prohibition enforcement agencies (see table 1). This information aggregates Prohibition enforcement operations from both federal and local authorities in most cases. I collected data on criminal judicial prosecutions from the Attorney General Annual reports, which are available at the Judicial District level only, for the years 1915-1936.

Demographic and Religious Data

For the first four decades of the Twentieth century, data on the distribution of religious ascriptions is available from the 1906, 1916, 1926, and 1936 decennial Censuses of Religions. The Censuses have comprehensive information about the number of adherents to each of the different faiths or churches in the United States.

Public Opinion Data

Most of the data comes from the state official rosters or "blue books", which states publish on an annual or biannual basis. The information for some of the states was found in the state archives, and for a few other referenda not reported in official sources, I took the data from local newspapers. A second major source of electoral data on the Prohibition issue are the election returns for the 21st Amendment Constitutional Convention elections, also found in the state rosters and some state archives.

Structural Estimation Data

The sample includes cities from all over the United States, and although the range of population sizes in this sample of cities goes from $51,000$ to 5.6 million (1920 numbers), admittedly this is a sample of urban communities. Of course, this is mainly due to the availability for the homicide rate data, which was reported on a population basis and for cities only. It is important to stress that the results should be seen as the effects of Prohibition in the most urbanized parts of the American society.

Table A5-1: Sample of Cities in the Structural Estimation

city	state	city	state	city	state
Akron	OH	Indianapolis	IN	Portland	OR
Albany	NY	Jersey City	NJ	Providence	RI
Atlanta	GA	Kansas City	KS	Reading	PA
Baltimore	MD	Kansas City	MO	Richmond	VA
Birmingham	AL	Los Angeles	CA	Rochester	NY
Boston	MA	Louisville	KY	Saint Louis	MO
Bridgeport	CT	Lowell	MA	Saint Paul	MN
Buffalo	NY	Memphis	TN	Salt Lake City	UT
Cambridge	MA	Milwaukee	WI	San Antonio	TX
Camden	NJ	Minneapolis	MN	San Francisco	CA
Chicago	IL	Nashville	TN	Scranton	PA
Cincinnati	OH	New Bedford	MA	Seattle	WA
Cleveland	OH	New Haven	CT	Spokane	WA
Columbus	OH	New Orleans	LA	Springfield	MA
Dallas	TX	New York	NY	Syracuse	NY
Dayton	OH	Newark	NJ	Toledo	OH
Denver	CO	Norfolk	VA	Trenton	NJ
Detroit	MI	Oakland	CA	Washington	DC
Fall River	MA	Omaha	NE	Wilmington	DE
Grand Rapids	MI	Paterson	NJ	Worcester	MA
Hartford	CT	Philadelphia	PA	Yonkers	NY
Houston	TX	Pittsburgh	PA	Youngstown	OH

Section 6.1 mentioned that in spite of being a dynamic model, Maximum Likelihood estimation was not subject to an initial conditions problem. The careful reader might have noticed that this requires the sample to cover years under no Prohibition and under Prohibition, while a few states were already

under Prohibition before 1911. Given the timing of the adoption of Prohibition across States (see figure 3), and the data availability, for Nashville and Memphis in Tennessee, Atlanta in Georgia, and Kansas City in Kansas, the sample covers Prohibition years exclusively. These three states officially adopted Prohibition in 1909, 1908 and 1880, respectively. Nevertheless, following the historical account on Prohibition in Tennessee, I code the cities in this state as being under Prohibition only starting in 1914. As mentioned in footnote 5.6.5, the governor of Tennessee decided not to enforce the constitutional amendment enacted in 1909, and Prohibiton only was enforced after the new Republican governor took office[52].

Although for Altanta, GA, and Kansas City, KS, the drunkenness arrests data also shows a fall only in 1917 (when War-time prohibition was adopted), suggesting little actual law enforcement of the state Prohibition laws (Atlanta's arrests fall from 18.4 to 12.2 between 1916 and 1917), there is no clear evidence that the laws were actually not being enforced. Instead of specifying a distribution for the unobserved homicide rate prior to 1911 for these three cities, I estimate the structural model assuming they enter Prohibition in 1917, and check the robustness of the model to excluding them from the estimation altoghether.

[52] The fact that Prohibition did not take place in Tennessee before 1914 can be corroborated directly by looking at the drunkenness arrests data. For example, this variable falls from 17.5 to 8.9 per $1,000$ people between 1913 and 1914 in Knoxville. Hilary House, Nashville's mayor at the time, even explicitly "acknowledged before the world that the state-wide Prohibition law is violated in Nashville... with knowledge and consent of the great majority of the people". Isaac (1965, p. 174)