NBER WORKING PAPER SERIES

INFANT MORTALITY AND THE REPEAL OF FEDERAL PROHIBITION

David S. Jacks
Krishna Pendakur
Hitoshi Shigeoka

Infant Mortality and the Repeal of Federal Prohibition
David S. Jacks, Krishna Pendakur, and Hitoshi Shigeoka
NBER Working Paper No. 23372
April 2017
JEL No. H73,I18,J1,N3

ABSTRACT

Exploiting a newly constructed dataset on county-level variation in prohibition status from 1933 to 1939, this paper asks two questions: what were the effects of the repeal of federal prohibition on infant mortality? And were there any significant externalities from the individual policy choices of counties and states on their neighbors? We find that dry counties with at least one wet neighbor saw baseline infant mortality increase by roughly 3% while wet counties themselves saw baseline infant mortality increase by roughly 2%. Cumulating across the six years from 1934 to 1939, our results indicate an excess of 13,665 infant deaths that could be attributable to the repeal of federal prohibition in 1933.

David S. Jacks
Department of Economics
Simon Fraser University
8888 University Drive
Burnaby, BC V5A 1S6
CANADA
and NBER
dsjacks@gmail.com

Krishna Pendakur
Department of Economics
Simon Fraser University
8888 University Drive
Burnaby, BC V5A 1S6
Canada
pendakur@sfu.ca

Hitoshi Shigeoka
Department of Economics
Simon Fraser University
8888 University Drive, WMC 4653
Barnaby, BC V5A 1S6
CANADA
and NBER
hitoshi_shigeoka@sfu.ca

1. Introduction

As is very well known, the United States from 1920 to 1933 embarked on one of the most ambitious policy interventions in the history of the modern nation state. Federal prohibition laws on the production, sale, and transportation of alcohol induced massive changes in the economic and social fabric of the then 48 states. And while contemporary prohibition movements gained traction across the world, nowhere were the impulses, proclivities, and traditions of such a large population subdued for so long. Naturally, given the scale of this intervention, prohibition has alternately been described as America's "noblest experiment" and its most ominous foray into social engineering.

Understanding the effects of federal prohibition is important with respect to a very sizeable historical literature on this topic (see Kyvig, 2000 and Okrent, 2010 among many others). And while this literature has advanced our understanding of the rise and fall of the prohibition movement as the confluence of specific political and social forces, there is surprisingly little research in assessing the economic and social outcomes of federal prohibition in the United States. In large part, this reflects a misunderstanding of the nature of prohibition. It was not in fact a monolithic policy change with national restrictions on alcohol "turning off" precisely in 1933. Instead, there was ample geographic and temporal heterogeneity in restrictions on alcohol after federal prohibition due to the decentralized nature of American government and the political concessions necessary to bring about repeal. In particular, the chief compromise for achieving the repeal of federal prohibition was in allowing for local option elections whereby local preferences determine whether a county, municipality, or even ward allows the sale of alcohol (Kyvig, 2000).

Exploiting a newly constructed dataset on county-level variation in prohibition status, this paper asks two questions: what were the effects of the repeal of federal prohibition—and thereby, potential maternal alcohol consumption—on infant mortality? And were there any significant externalities from the individual policy choices of counties and states on their neighbors? Our focus on infant mortality stems from the fact that it is not only a key determinant of life expectancy but also a rough indicator of population health. What is more, infant mortality represents an acute, rather than a chronic, outcome of potential alcohol consumption, making identification a slightly easier, but still challenging task. There is a substantial literature in understanding the drivers of infant mortality in a historical context (cf. Alsan and Goldin, 2015; Clay *et al*., 2016; Cutler and Miller, 2005; Fishback *et al*., 2001, 2007, 2011). However, to our knowledge, this is the only study that considers the effects of the repeal of federal prohibition on infant mortality or—for that matter—any other outcome variable.

An important methodological contribution of the paper then comes in explicitly recognizing the possibility of policy externalities across county borders. Thus, after repeal of federal prohibition, it is not only an individual county's choice of prohibition status which matters but also the prohibition status of its neighbors. In this manner, we distinguish among counties which opt to allow the sale of alcohol within their borders ("wet" counties), counties which opt to continue with alcohol prohibition and find themselves with neighbors which do the same ("bone dry" counties), and—critically—counties which opt to continue with alcohol prohibition but find themselves with a wet neighbor ("dryish" counties).

A further methodological contribution is to take the count nature of the data seriously in our empirical model, potentially improving on existing methods. Much of the literature on the causes of infant death uses OLS regression modeling to explain variation in (logged) infant

3

mortality rates (c.f., Anand and Bärnighausen, 2004 and Baird *et al*., 2011). However, if either infant mortality rates or the numbers of births are very low, the observed infant mortality rates become highly discrete, e.g., with two births, infant mortality rates of only 0.0, 0.5, and 1.0 can be observed. In such instances, OLS may be inappropriate. We implement a binomial fixed-effect model that takes into account the facts that infant death is a relatively low incidence phenomena and that many counties have a quite small number of births in any given year.

Using this approach, we find that wet status raised infant mortality by roughly 2%, or about 1.2 additional infant deaths per 1000 live births. Allowing for potential policy externalities from neighboring counties turns out to be important as well: we find that dryish status raised baseline infant mortality by roughly 3%, or 1.77 additional infant deaths per 1000 live births. Cumulating across all affected counties and over the years 1933 to 1939, these results imply a minimum of 13,665 excess infant deaths that could be attributed to the repeal of federal prohibition in 1933.

Our paper is broadly related to a literature which focuses on assessing the effects of state-level measures prior to federal prohibition on variables such as the incidence of adult heights (Evans *et al*., 2016), cirrhosis (Dills and Miron, 2004), and homicide (Bodenhorn, 2016). However, we are the first to study the effects of federal prohibition's repeal and the first to study prohibition in the context of *county-* as opposed to *state*-level variation in prohibition laws. Here, we argue that a priori county-level information is likely more meaningful, and below, we also directly address the issue of endogeneity that these changes in prohibition status necessarily entailed. Our paper is also related to recent work by García-Jimeno (2016) which considers the effects of federal prohibition on city-level crime during the period from 1920 to 1933. Here, local enforcement of federal prohibition laws not only generates extra-judicial homicides and

4

other forms of crime but also responds endogenously to perceptions of its efficacy in the immediate past. Thus, our paper shares at least one element with his work, namely an appreciation of the potential divergence between *de facto* and *de jure* prohibition status, both during and after federal prohibition. This is a point which we return to below and which flavors the interpretation of our results.

Apart from historical interest, understanding the effects of federal prohibition is important with respect to contemporary policy issues related to alcohol and the control of illicit substances. First, this particular historical setting has unique advantages in estimating the effect of restrictions on alcohol on infant mortality. The US Surgeon General's initial warning about the risks associated with alcohol consumption during pregnancy was only issued in 1981. Thus, the general public at the time had little knowledge of the potential negative effects of alcohol consumption during pregnancy on child development. Thus, our estimates are not confounded by differences in avoidance behaviors—both avoiding conception and avoiding drinking—by mothers of different socioeconomic status (Nilsson, 2016).

Second, recent studies on the effects of alcohol restrictions have predominately focused on relatively small differences in variables such as the minimum drinking age or the availability of beer as opposed to spirits (Barreca and Page, 2015; Carpenter and Dobkin, 2009; Nilsson, 2016). However, little is known about the effects of more stark policy changes where the relative price of alcohol is more dramatically altered. We note that the scope for policy interventions is still large: although information about the risks associated with alcohol consumption during pregnancy is now widely understood in the US, over 50% of women of childbearing age drink while over 10% of women continue to drink during pregnancy (Tan *et al*., 2015).

Finally, our paper speaks to a related literature in public economics which considers differential taxation across state borders in the presence of competition in local markets and its effects on firm pass-through and, thereby, consumer prices for items like alcohol and cigarettes (cf. Doyle and Samphantharak, 2008; Harding *et al.*, 2012; and Lovenheim, 2008). This is particularly true if we conceive of prohibition and its repeal as having vitally affected the price— and not necessarily the availability—of alcohol. However, to our knowledge, few papers have addressed the issue of policy externalities, or how one location's policy choice affects outcomes in another, which this paper so strongly emphasizes. Lovenheim and Slemrod (2010) is a notable exception in that it finds that one state's reduction in the minimum legal drinking age could lead to a substantial increase in teenage traffic fatalities in neighboring states. Likewise, Johansson *et al.* (2014) find higher rates of workplace absenteeism in Sweden after a cut in Finnish alcohol taxes. What differentiates our work in this respect is the focus on an entire country (the United States) at the lowest level of geographic aggregation possible (counties) for a more dramatic change in alcohol policy (the repeal of federal prohibition).

This paper also provides at least one valuable history lesson for the present-day debate on legalization of illicit substances, in particular, the recent spate of state-level legislation related to marijuana. A key insight of our paper is that infant mortality in this period was not solely driven by any individual county's choice of prohibition status but rather what its neighboring counties' choice of prohibition status was. That is, a county's choice to go wet and allow for the sale of alcohol in its borders vitally affected infant mortality in neighboring counties which chose to remain dry. It is telling that in every historical discussion on the relative merits and demerits of county-level repeal known to us, none make reference to the possibility of one county's choice affecting another. Likewise, the debate on the relative merits and demerits of state-level

6

legalization of marijuana has failed to adequately address the possibility of cross-jurisdictional externalities such as the one documented in this paper.

The rest of the paper proceeds as follows. Section 2 lays out the historical context and the relationship between infant mortality and the repeal of federal prohibition, and provides a simple framework for thinking about local alcohol consumption, counties' choice of whether or not to allow for alcohol sales, and the effects of doing so. Section 3 introduces the underlying data while Section 4 introduces our empirical model. Section 5 presents our results on infant mortality, and Section 6 concludes by considering caveats to our study and avenues for future research.

2. Context

In this section, we lay out the historical context surrounding the rise and fall of federal prohibition, highlighting those institutional features of its repeal which are most bearing for our analysis. Then, we lay out a simple framework for how preferences for alcohol consumption at the local level likely determine both prohibition status and the size of the related treatment effects.

2.1 Historical background

On a rising tide of an anti-alcohol movement led by rural Protestants and social progressives, the US Senate proposed a constitutional amendment to affect a federal prohibition on alcohol on December 18, 1917. With the approval of 36 states by January 16, 1919, the 18th amendment was, thereby, ratified with the country becoming dry on January 17, 1920. This entailed a near-complete prohibition on the production, sale, and transportation of alcohol. But

7

by no means did this entail the complete unavailability of alcohol as the individual possession and consumption of alcohol was not explicitly prohibited, allowing for wide differences in legislation and enforcement along these lines at the municipal, county, and state level. Rather, prohibition is best thought as having substantially increased the price of alcohol (Cook, 2007).

Surprisingly large effects on quantities were also forthcoming: total alcohol consumption fell by as much as 70% in the 1920s (Miron and Zwiebel, 1991) with per-capita consumption only reaching half of its pre-prohibition peak throughout the 1930s and only surpassing this level in the 1970s (Blocker, 2006). However, initial wide-spread support for federal prohibition was eroded throughout the 1920s in the wake of rising criminal activity, doubts over prohibition's efficacy, and concerns over the reach of the federal government which it necessarily entailed (Okrent, 2010).

Turning to its demise, the proverbial nail in the coffin for federal prohibition arose from the fiscal straits of the Great Depression. Prior to 1920, 15% of government revenues came from alcohol with the federal government alone collecting nearly $500 million in 1919 (Blocker, 2006), or nearly $7 billion in 2015 dollars. Thus, starved of other sources of funding, various levels of government increasingly viewed the sales of alcohol as a potential source of revenue. The opening salvo in repealing federal prohibition came on March 22, 1933, when Franklin Roosevelt amended the Volstead Act (or National Prohibition Act), allowing for the resumption of low-alcohol beer consumption and production (Okrent, 2010). From there, popular and political support for prohibition quickly eroded, and the 18th Amendment was repealed on December 5, 1933, with ratification of the 21st Amendment to the US Constitution.

However, the process of repeal was decidedly—and deliberately—not uniform. The chief compromise for achieving ratification of the 21st Amendment was in allowing for local option

elections to determine liquor laws deemed appropriate for local conditions (Kyvig, 2000). These

elections give the electorate the right to vote on liquor control by referendum. That is, local

preferences determine whether a county or municipality prohibits the sale of alcohol. At the

same time, many states opted out from local option elections entirely while others allowed for

referenda to be held at the state-, county-, city-, or even ward-level. The transition from

prohibition was, in many instances, very rapid: by 1935, 2,120 counties became wet in some

form while 991 counties stayed dry (Strumpf and Oberholzer-Gee, 2002).[1]

With respect to infant mortality, we draw on a large body of work which explores its

causes in a historical context (cf. Alsan and Goldin, 2015; Clay *et al*., 2016; Cutler and Miller,

2005; Fishback *et al*., 2001, 2007, 2011). We note, however, that this work has little to say about

the mechanisms by which the repeal of federal prohibition could have influenced infant

mortality. Instead, we turn to a substantial medical literature linking maternal alcohol

consumption and both compromised infant immune systems and reduced birth weight—two key

determinants of subsequent infant death (cf. Mills *et al*., 1984; and Olegård *et al*., 1979;

Strandberg-Larsen *et al*., 2009).

Unfortunately, we lack any information of maternal alcohol consumption at the

individual or aggregate level for this period.[2] Having no other more plausible prior, our proposed

causal mechanism for this paper runs from the repeal of federal prohibition to potential maternal

alcohol consumption and from there to infant mortality. At the same time, we acknowledge that

---

[1] More precisely, of the 2,120 counties that became wet in some form, 341 counties were of mixed status—that is, a wet county with at least one dry municipality or vice-versa. In what follows, we treat mixed counties as equivalent to wet counties as our main results are unchanged when making this distinction (results available by request) and are omitted here for expositional purposes.

[2] To our knowledge, other proxies for alcohol consumption like the number of retail outlets for alcohol, retail sales of alcohol, or tax revenues from alcohol sales are not systematically available at the county level. Likewise, average birth weight and the general health of newborns was not recorded for this period.

other forces may have been at work such as paternal alcohol consumption and its effects on

domestic violence, the general home environment, household budgets, and/or pre-natal

investment. We necessarily leave this task for future work, citing a lack of relevant data at the

present.

2.2 A simple framework for analysis

In this brief subsection, we provide a simple framework demonstrating how the

distribution of the willingness-to-pay (WTP) for alcohol within a county might *simultaneously*

determine whether or not a county goes wet *and* the treatment effect of that choice (please refer

to Appendix A for more detail). We motivate the framework as follows. For individuals, assume

that consuming alcohol is a binary decision. They do so if their WTP for alcohol exceeds the

prevailing price of alcohol.

We assume that each county can be characterized by its distribution of individual WTP

for alcohol and that this distribution differs by county. To simplify, we also assume that the

distribution of WTP is symmetric and unimodal so that the median voter is at the top (mode) of

the density function. One such distribution is shown in Figure 1a depicts the distribution of WTP

in a county where people generally have a high WTP. Under prohibition, alcohol sales are

illegal, making it difficult to purchase and, thereby, expensive.[3] The effective price of purchasing

and consuming alcohol includes the wedges associated with this difficulty and illegality and is

denoted in Figure 1a by the line labeled "dry price". The area to the right of the dry price gives

the fraction of the population of the county that consumes alcohol when the county's *de jure*

---

[3] Again, Cook (2007) among others is clear that prohibition never entailed a complete lack of availability of alcohol
in affected counties, rather it is best thought of as having raised the price of alcohol by a factor of three to six.

status is dry. For this county, where generally people have a high WTP, a large fraction of people consume alcohol even though it is prohibited.

The price of purchasing and consuming alcohol is lower when a county becomes wet as the sale of alcohol may be done more easily and openly. The "wet price" is shown by the vertical line to the left of the dry price. The *treatment effect* of this price change is to switch people whose WTP lies between the wet price and the dry price from non-consumers into consumers of alcohol. In Figure 1a, the treatment effect of a county changing its prohibition status—that is, the average treatment effect on the treated (ATT) of becoming wet—is equal to the shaded area under the curve between the wet and dry prices. Thus, we see that for this county where the WTP is generally high, the magnitude of the treatment effect is small because many people were already consuming alcohol prior to repeal.

Consider whether or not this county will choose to become wet. Any individual whose WTP exceeds the wet price will gain from the reduction in the cost of alcohol from the dry to wet price. The WTP distribution shown in Figure 1a has the feature that its median (located at the top of the curve) is to the right of the wet price, so the majority of people would prefer to that their county become wet. Naturally, in a strict majority-rule setting, this county would choose wet status (or, more generally, it would have a higher probability of choosing wet status).

Likewise, if we consider a county where the distribution of WTP was shifted to the right, so that people liked alcohol even more, all these features would be amplified. The level of alcohol consumption under prohibition would be higher, the treatment effect of becoming wet would be smaller, and the probability of becoming wet would be higher. Thus, we have that the distribution of WTP determines both the probability of treatment and the treatment effect.

If WTP distributions—that is, preferences—vary across counties but are fixed over relatively short periods of time (as in our sample), then the inclusion of county fixed-effects would account for all the drivers of both the probability of treatment and the treatment effect. Therefore, the inclusion of county fixed effects would provide us with "selection-on-observables" (or, in other words, exogeneity of treatment). Furthermore, since the distribution of WTP differs across counties, we would have heterogeneous treatment effects. Taken together, this implies that fixed-effect models should obtain unbiased estimates of the average treatment effect on the treated for those counties choosing wet status.

Now, consider a county where generally people have a low WTP as shown in Figure 1b. In this county, only a small fraction of the density lies to the right of the dry price, so that only a small fraction of people chooses to consume alcohol under prohibition. The treatment effect of becoming wet, given by the region under the curve between the wet and dry prices, is quite large. However, unlike the county shown in Figure 1a, the median individual in this county has a WTP lower than the wet price. That is, the majority of people in this county have nothing to gain from their county becoming wet as they are non-consumers, regardless of their county's *de jure* prohibition status. Therefore, in a strict majoritarian setting, this county would not choose to end Prohibition, or, more generally, has a lower probability of becoming wet than the county depicted in Figure 1a.

Finally, consider a county that chooses to stay dry but which has a wet neighbor. Thus, it could face a change in the price of alcohol even though it did not choose to become wet itself, but instead, became dryish due to the choice of its neighbor. As in wet counties, the "dryish price" in Figure 1b is less than the dry price because alcohol may now be legally purchased in the neighboring county. But there are two frictions that make the dryish price strictly larger than

the wet price: (1) the driving distance (or other transportation cost) to the neighboring county line reduces the effective magnitude of the price drop; and (2) the prospect of enforcement of local prohibition within the county likely also reduces the effective magnitude of the price drop.

In Figure 1b, the treatment effect of staying dry but having a neighbor which became wet—that is, the average treatment effect on the treated (ATT) of becoming dryish—is equal to the shaded area under the curve between the dryish and dry prices. Here, the treatment is a smaller price change than going all the way to the wet price.[4] Again, with the inclusion of county fixed effects in our empirical model, we should be able to obtain unbiased estimates of the ATT of the dryish treatment. Finally, by comparing the size of the shaded areas between Figure 1a and Figure 1b, we see that the ATT for becoming wet *could plausibly be smaller* than the ATT for becoming dryish.

There are several lessons that we draw from this framework that will be important for understanding our empirical results: (1) the distribution of WTP varies across counties and determines whether or not a county goes wet *as well as* the treatment effect of that choice—that is, we will have heterogeneous treatment effects; (2) if the WTP distributions—or in other words, the underlying preferences of individuals—within counties are invariant over relatively short periods of time (as in our sample), then the inclusion of county fixed effects in the empirical model is sufficient to obtain unbiased estimates of the ATT of becoming wet or becoming dryish;[5] and (3) the ATT for becoming wet could plausibly be smaller than the ATT for becoming dryish.

---

[4] This indicates that the dryish treatment is a smaller treatment than the treatment which would obtain if a county that wished to stay dry was *forced* to go wet. In other words, our dryish estimate then serves as a conservative estimate of the *hypothetical* treatment of becoming wet in a county that would otherwise choose to stay dry.
[5] In Section 4 below, we more fully consider the potential bias induced if we omit variables correlated with treatment that are also correlated with infant mortality and the means at our disposal to deal with this issue.

13

3. Data

Our data are drawn from three main sources: annual, county-level infant mortality and live births have been extracted from the *Vital Statistics of the United States*; annual, indicators of county-level prohibition status have been constructed from Strumpf and Oberholzer (2002) and contemporary sources; and other county-level covariates are available from the US Census.

3.1 Infant mortality

Annual counts of infant deaths and live births from 1933 to 1939 for the 3,000+ counties of the continental United States are available from Fishback *et al*. (2011). The choice of a start date in 1933 is predicated by the fact that mortality statistics for Texas, with its 254 counties, only begins in this year. The choice of an end date in 1939 is predicated by the fact that the vast majority of changes in prohibition status had occurred by 1938. We also wish to avoid any confounding effects of the mobilization effort for World War II. We do, however, extend the sample in a robustness exercise below.

To reiterate, we expect increases in the availability of alcohol across counties should be associated with higher infant mortality rates, given the documented role of alcohol in both compromising infant immune systems and inducing low birth weights (cf. Mills *et al*., 1984; and Olegård *et al*., 1979; Strandberg-Larsen *et al*., 2009). Figure 2 depicts infant mortality rates by prohibition status for every year from 1934 to 1939, weighted by the number of births in a county. It is important to note that these series do not hold constant the composition of counties under the various headings, so there may be some role for changes in prohibition status in driving the underlying trends.

14

On average, the infant mortality rate drops from 60.0 per thousand live births in 1934 to 50.2 per thousand in 1939. However, there seem to be significant differences in levels and trends across the three categories of bone dry, dryish, and wet. To begin, the highest infant mortality rates in 1934 are for bone dry counties, then dryish counties, and then finally wet counties. Perhaps more tellingly for our purposes is not the difference in initial levels, but instead the difference in subsequent trends. Bone dry counties register the steepest proportional decline in infant mortality rates over this period, moving from the highest levels to the second lowest. Compare this performance to that of wet counties which register the least steep proportional decline in infant mortality rates while maintaining the lowest levels and to that of dryish counties which register nearly the same proportional rate of decline as wet counties, moving from the second highest levels of infant mortality rates to the highest. Compositional issues may explain part of this performance, but as we show below, dryish status is systematically associated with higher infant mortality rates with a wide set of controls and under a wide range of specifications.

3.2 County-level prohibition status

Ideally, we would like county-level information on alcohol consumption, particularly for pregnant women or, at least, women of child-bearing age. Of course, this type of data is not available before, especially during, or even after prohibition. Another possibility would be to rely on other legal restrictions on alcohol, yet liquor laws in the United States appear in very diverse forms: among other things, individual counties and states continue to limit the maximum alcohol content of specific types of beverages sold within their borders, specify whether alcohol can be sold for off- or on-premise consumption for specific types of establishments, and/or place restrictions on the day and time of alcohol sales (so-called "blue laws").

15

Instead, we rely on the sharpest distinction in prohibition status possible: dry versus wet. That is, we seek to compare outcomes for those counties for which no sales of alcohol are permitted (dry) to those for which at least some sales are permitted (wet). Again, we also make the critical distinction in between those counties which are dry and have no wet neighbors versus those counties which are dry and have at least one wet neighbor. Thereby, we decompose all dry counties into either bone dry or dryish counties, respectively.

To achieve this goal, we build on previous data collection efforts. Our starting point is in reconstructing the prohibition status of counties in 1935 and 1940, depicted in the maps of Strumpf and Oberholzer (2002). We then supplement these with new sources to fill in the gaps (Culver and Thomas, 1940; The Distilled Spirits Council, 1935, 1941; Harrison, 1938; Thomas and Culver, 1940). For a small number of counties, it was required to use LexisNexis to establish the year in which there was a change in their prohibition status. Thus, we make a significant contribution with respect to data by reconstructing the prohibition status of all continental US counties for the critical post-repeal period from 1934 to 1939.[6]

Figure 3 depicts the proportion of counties by prohibition status for the longer period from 1930 to 1942. We assume all counties are bone dry from 1930 to 1933 (in our results below, we partially relax this assumption by excluding those counties on the Canadian or Mexican border in 1933). By 1939, this proportion had dropped from 100% to slightly below 15%. Likewise, we observe the proportion of wet counties rising from 0% in 1933 to slightly above 70% in 1939 and the proportion of dryish counties rising from 0% to slightly above 15% in 1939. Thereafter, there is very little aggregate change in prohibition status throughout the

---

[6] In a larger project on the contemporaneous and long-run effects of prohibition, we have reconstructed the status of all continental US counties from 1885 to 1984.

1940s. Given that the vast majority of changes in prohibition status occurred by 1938, we estimate our empirical model for the period from 1933 to 1939 as it represents the minimal dataset for identifying the effects of repeal. That is, adding years prior to 1933 and after 1939 adds very little by way of variation in our independent variable of interest, namely individual counties' prohibition status. This relatively short panel is also beneficial in that we believe that preferences for alcohol are unlikely to have changed very much over such a short period of time as discussed in Section 2.2 above.

Figure 4 depicts the spatial distribution of prohibition status by year from 1933 to 1939. There, it is clear that by 1935 the remaining hold-out states for prohibition were along the central axis of the US (Kansas, North Dakota, and Oklahoma) along with large parts of the Southeast (Alabama, Georgia, Mississippi, and Tennessee). This constellation changed considerably by 1937 with Alabama and North Dakota jettisoning state-wide prohibitions and opting for local option. This along with the steady change in prohibition status at the county level for Georgia, Tennessee, and Texas in later years makes for what we hope is ample variation in our dryish measure.

3.3 Additional covariates

In determining a valid specification relating infant mortality rates to changes in prohibition status, Figure 2 suggests a potentially large role for both a common trend and time-invariant county characteristics. Thus, we include county and time fixed effects in all of our specifications. There is also a large historical literature delineating characteristics which shaped support for prohibition and which may be useful as further controls in explaining variation in infant mortality rates. This literature points to strong preferences for dry status before and during

17

the period of repeal among the native-born, Protestants, rural inhabitants, and women (Okrent, 2010). To this list, we also include information at the county level on the proportion of blacks, the number of medical institutions per 1000 people, the number of hospital beds per 1000 childbearing age women, per capita New Deal spending, per capita retail sales as a proxy for income, population, and the unemployment rate as previous research has indicated that these variables influenced infant mortality in this period (Fishback *et al*., 2001, 2007).[7]

Table 1 provides the definition and sources of our control variables while Table 2 reports the mean of the infant mortality rate in 1933 (both in levels and logs) along with the means of our proposed (predetermined) county-level control variables in or around 1933. These are reported for the full sample in column (1) and for three mutually exclusive categories: always dry, stay dryish, and ever wet. *Always dry* counties stayed bone dry during our sample period from 1933 to 1939. *Stay dryish* are the counties that became dryish at any time from 1933 to 1939 and stay dryish until 1939. *Ever wet* counties became wet at any time from 1933 to 1939. As such, counties that became dryish first but eventually became wet are included in *ever wet* and not in *stay dryish*.

Table 3 reports the results of hazard analysis to determine whether predetermined county characteristics have any predictive power for the timing of changes in prohibition status. Columns (1) and (2) concern the probability of counties voluntarily changing to wet status with the former incorporating our baseline controls and the latter incorporating our full set of potential controls. We distinguish between these two set of controls in that New Deal spending per capita, the number of medical institutions per capita, and the number of hospital beds per capita could

---

[7] We use these variables strictly as controls. They would not make good instruments for IV regression analysis because they all plausibly have direct effects on infant mortality, and so would not satisfy the exclusion restriction.

arguably be endogenous if, say, wet status afforded greater tax revenues and a correspondingly greater expenditure on medical facilities at the county level.

In relation to our baseline controls, we can see that all of the variables are statistically significant in explaining the timing of counties voluntarily changing to wet status and match with nearly all of our priors. Counties with higher levels of incomes, higher levels of population, higher levels of unemployment, proportionally more immigrants, proportionally more non-Protestants, and proportionally more whites are more likely to choose wet status while counties with proportionally more women of child bearing age are less likely to choose wet status. There are also some statistically significant results which run counter to our expectations, namely that counties with proportionally more women of all ages are more likely to choose wet status while counties which are more urban are less likely to choose wet status. Column (2) replicates these results and also suggests that countries with higher per capita New Deal spending and a higher per capita number of medical institutions are more likely to choose wet status. Cumulatively, these results suggest that the historical literature on the drivers of the onset of prohibition do indeed capture some of the drivers of the end of prohibition at the county level. They also suggest that the inclusion of county fixed-effects is likely important to ensure that omitted variables do not bias our estimates of the ATT for a county becoming wet.

Columns (3) and (4) show the shifters to the hazard function for counties involuntarily changing to dryish status. Of the included control variables only the proportion of women of childbearing age and retail sales per capita—again, a proxy for income—register as statistically significant. Additionally, most of the estimated magnitudes in columns (3) and (4) are very small. Even if we had four times as many dryish observations yielding twice the precision, the regressors would still be statistically insignificant. These results suggest that dryish counties are

essentially a random sample of all dry counties, so that we may interpret our dryish estimate as an unbiased estimate of the observed effect of a dry county being treated with a wet neighbor.

The bottom line from the hazard analysis is as follows. The estimated hazard functions suggest that the date of transition to wet status is highly predictable from observed time-invariant county characteristics. Standard difference-in-difference reasoning tells us that if potential endogeneity is driven by time-invariant but heterogeneous factors like alcohol preferences across counties, then the inclusion of fixed effects in our models will yield unbiased estimates of the treatment effect of a county becoming wet.[8] Naturally, the role of unobservables in driving changes in both dryish and wet status may still be of concern. To account for this, we interact all county-level control variables in Table 3 with linear time trends in our specifications as in Acemoglu et al. (2004) and Hoynes and Schanzenbach (2010) to control for differences in trends across counties which may be spuriously correlated with counties' prohibition status.[9]

Difference-in-difference estimation also employs a common-trends assumption to identify treatment effects. Under the assumption that treated counties would have followed the same time trend as untreated counties had they not been treated, the difference in rates of decline or growth between treated and untreated counties equals the true treatment effect. One way to gauge the validity of this assumption is to compare the time trend *before any treatments occur* of counties that are eventually treated with those that are never treated (sometimes called a "pre-trend"). If they are parallel, then the common-trends assumption is satisfied before any treatments occur and, therefore, might well be satisfied during treatment periods.

---

[8] Thus, our estimates of the treatment effects for wet may be unbiased, but in the absence of instruments to support IV methods, we are unable to provide additional evidence for exogeneity (such as a Hausman test). In contrast, for dryish status, we see little evidence that even county-level time-invariant factors matter for predicting the date of transition to dryish.

[9] The only exceptions in this regard are access to medical institutions, the number of hospital beds, and per-capita retail sales which are time-varying and not predetermined as is the case for our other control variables.

Figure 5 tracks infant mortality rates (weighted by the number of births in a county) for the period from 1928 to 1933. Here, we use only the 2,670 counties where we observe vital statistics back to 1928, which leaves out, for example, Texas. We consider pre-trends back only to 1928 because the number of states reporting vital statistics drops drastically before that. We employ the same classification scheme as in Table 2: three mutually exclusive categories of *always dry*, *stay dryish*, and *ever wet*. Figure 5 holds the composition of counties constant, but only does so by ignoring the timing of counties' changes in prohibition status (whereas Figure 2 incorporated the timing of counties' changes in prohibition status, but only did so by not holding the composition of counties constant).

A general decline in infant mortality rates is evident throughout this period for all three county types, and the general ordering of counties by type is preserved when considering 1928 and 1933 in isolation: infant mortality rates are increasing from counties that were *ever wet* to counties that were *always dry* to counties that *stay dryish* throughout. In the intermediate years, *always dry* and *stay dryish* are strikingly parallel, but *ever wet*—while still downward sloping— does not seem exactly parallel to the other two types. Thus, the pre-trends exhibit similar overall trends for all three types (that is, decreasing infant mortality for all three types), but there may be some room for doubt in regard to the common-trends assumption for *ever wet* counties due to their dissimilarity with the other two types between 1929 and 1932.

4. Empirical model

The most straightforward empirical model in this regard would likely be to regress the (logged) infant mortality rate on prohibition status with county and year fixed effects. But as an empirical fact, many counties have small populations with consequently a small number of births

21

and, therefore, a very small number of deaths. For instance, the median number of infant deaths in a county-year in our sample is 15, and a quarter of county-years have less than 10 infant deaths. Thus, it is desirable *a priori* to not treat the outcome as continuous and, instead, take the count nature of the data seriously in our estimation.

More formally, we investigate how alcohol prohibition (the "treatment") affects the incidence of infant mortality (the "response"). Our response variable is the number of infant deaths, *D*, in a county-year. Given the number of births, *B*, the infant mortality rate, *I*, is given by *I=D/B*. Our treatment variables are county-year level indicators of whether or not a county is itself wet, *W*, and of whether or not a county has at least one neighboring county that is wet, *N*. These are coded as to be mutually exclusive by giving priority to *W*; for example, a county that is wet and has a wet neighbor has *W=1* and *N=0*. The excluded category is a bone dry county that has no wet neighboring counties and is itself not wet, thus, having *W=N=0*.

For all prohibition treatment and response variables, we have panel data on counties *c=1,…,C* and time periods *t=1,…,T*. Our source on infant mortality, the *Vital Statistics of the United States*, counts infant deaths in the year after birth, and births occur roughly nine months after conception. Thus, most infant deaths caused by the relaxation of alcohol prohibition would occur in the years following the change in legal status and not during that year. What is more, we lack complete information on when changes in prohibition status were implemented within any given year for any given county. Consequently, we set W_{ct} and N_{ct} equal to one in all the years *following* the change in legal status and equal to zero in the year of and all years preceding the change in legal status. We additionally include variables allowing for partial treatment effects in the year of change in legal status. The regressors W_{ct*} and N_{ct*} are equal to one in the year of the

legal change and equal to zero in all other years. Let P_{ct} be the vector of prohibition status variables $[W_{ct}, N_{ct}, W_{ct*}, N_{ct*}]$.

Infant death is a dichotomous outcome (that is, analogous to a coin toss) for a given birth. In our data, we observe totals of births and deaths in a given county-year, but these are in fact the result of summing up dichotomous infant deaths. The correct statistical model for this data environment is, thus, the binomial distribution, which is the distribution of sums of Bernouilli trials. Assume that each birth b_{ict}, for $i=1,...,B_{ct}$, in county c in time period t is an independent Bernoulli trial with a probability γ_{ct} that the birth results in an infant death. There are a total of B_{ct} births in a given county-year, and our measured outcome is the total number of infant deaths D_{ct} in that county year. The Bernoulli structure of infant deaths implies that the probability mass function of the number of infant deaths, D_{ct}, given the number of births, B_{ct}, follows the binomial distribution, denoted $Bin(D_{ct}, B_{ct}, \gamma_{ct})$, and given by

$$[1] \quad Bin(D_{ct}, B_{ct}, \gamma_{ct}) = \frac{B_{ct}!}{D_{ct}!(B_{ct} - D_{ct})!} (\gamma_{ct})^{D_{ct}} (1 - \gamma_{ct})^{B_{ct} - D_{ct}}$$

The probability γ_{ct} is our object of interest, and it depends on our prohibition status variables, P_{ct}, and on observed control variables, X_{ct}. It is quite reasonable to think that unobserved characteristics of counties could both cause a county to remain dry and influence its infant mortality. Thus, any sensible model should have county-fixed effects, denoted θ_c. Finally, to account for national time trends in infant mortality, we include time dummies, δ_t.

Since probabilities lie in between zero and one, and since for infant mortality, the probabilities cannot be exactly equal to zero, we specify the probabilities as:

$$[2] \quad \gamma_{ct} = \frac{\exp(P_{ct}\alpha + X_{ct}\beta + \theta_c + \delta_t)}{1 + \exp(P_{ct}\alpha + X_{ct}\beta + \theta_c + \delta_t)}$$

Again, it may be tempting to regress the infant mortality rate in a county-year, I_{ct}, on prohibition

treatment variables, county and year fixed-effects, and other time-varying control variables.

However, this could be misleading because many counties have a quite small number of births

Since infant death is already a low incidence phenomena, linear regression could easily yield

predicted values for the infant mortality rate of less than zero. If it does, then the OLS estimator

is inconsistent (Horrace and Oaxaca, 2006). This problem is not solved, for example, by using

logged infant mortality rates because in that case, county-years with zero infant deaths have to be

dropped or scaled by adding an arbitrary constant to all observations. Both of these strategies

induce inconsistency in the estimator.

We instead estimate the model by maximum likelihood. Consequently, we deal directly

with the count nature of the data and avoid the inconsistency of the OLS estimator when the

incidence of infant mortality γ_{ct} is low. The MLE for this fixed-effect binomial model is given by

$$[3] \quad \max_{\alpha,\beta,\gamma,\delta} \sum_{t=1}^{T}\sum_{c=1}^{C} \ln Bin(D_{ct}, B_{ct}, \gamma_{ct}(\alpha, \beta, \gamma, \delta, P_{ct}, X_{ct}))$$

where γ_{ct} is given by [2] as above. We provide consistent estimates of parameters following

Hahn and Newey (2004).[10]

---

[10] The MLE for the fixed effects binomial model suffers from an incidental parameters problem which may induce bias in the MLE when T is fixed (and small) (Machado, 2004). Incidental parameters problems arise in fixed-T nonlinear panel models if the presence of fixed effects (in our case, for counties) induces bias in the estimated treatment effects (in our case, the coefficients on prohibition status). In panel models where T goes to infinity, there are typically no incidental parameters problems because each fixed effect may be estimated consistently. In linear panel models, we can typically difference the model so that the fixed effects do not need to be estimated. In nonlinear fixed-T settings, the incidental parameters problem is roughly that the inability to consistently estimate or difference out fixed effects induces bias in estimated treatment effects. In Monte Carlo experiments, Machado (2004) finds that the incidental parameters bias in the MLE is small for $T > 4$. Hahn and Newey (2004) provide an analytical bias-correction for general nonlinear fixed effects models for the case where T is increasing much more slowly than C. In the case of our binomial model, this bias correction may be written explicitly in terms of observed variables and is straightforward to compute. In the main text of this paper, we present only bias-corrected MLEs. Consistent with Machado's observation that bias is small when T is not small, none of our bias corrections exceed 1% of the uncorrected MLE.

Note that since we have a full vector of county dummies, there is no intercept term in this model for the probabilities in equation [2]. Note also that this estimator does not include the observed infant mortality rate, I_{ct}. Instead, it maximizes the likelihood given the probability of observing each possible value of D_{ct} for a given B_{ct} when the probability of death for each birth is γ_{ct}. This model, thus, takes into account the heteroscedasticity implied by the number of births varying across county-years and accounts for the count nature of the dependent variable D_{ct}. Consequently, the MLE has an efficiency advantage in comparison to OLS, analogous to the efficiency gain from weighted least squares when using grouped data.

The marginal effect of changing prohibition status P_{ct} on the probability of infant mortality is equal to $\alpha\gamma_{ct}(1-\gamma_{ct})$. In the case where infant mortality rates are low (e.g., they average around 5% in our sample), this is approximately equal to $\alpha\gamma_{ct}$, so we can interpret the estimated treatment effect (α) as approximately equal to the semi-elasticity of the probability of infant death with respect to treatment. Another model where estimated coefficients are approximately equal to the semi-elasticity of the probability of infant death with respect to treatment is the regression of the logged infant mortality rate, $\ln(I_{ct})$, on P_{ct}, X_{ct}, and county and year fixed effects. Linear regression estimators are very common in the literature on infant mortality. For instance, Alsan and Goldin (2015) use such a specification to estimate the effects of clean water and sewerage on infant mortality in the greater Boston area from 1880 to 1915 while Clay *et al.* (2016) do the same in estimating the effects of pollution from burning coal for the entirety of the United States from 1938 to 1962. Consequently, we present results from this type of regression for easy comparability with previous research in Appendix B. There, the most important message is that WLS—but not OLS—estimates of our models yield very similar results to MLE.

A final consideration is the potential bias induced if we omit variables correlated with treatment that are also correlated with infant mortality. As described in section 2.2, we believe that the choice to become wet is necessarily related to the county-level distribution of preferences for alcohol consumption. To the extent that such preferences are fixed over time, the inclusion of county fixed-effects fully accounts for such preference variation and, thus, yields unbiased estimates for both dryish and wet. Below, our focus is on the years from 1933 to 1939, and we believe that preferences are unlikely to have changed very much over such a short period of time. However, if there are county-level, time-varying factors that are correlated with treatment and infant mortality, then our estimates may suffer from bias. One possibility here is learning about the consequences of alcohol consumption/enforcement as other counties become wet (similar to the argument of García-Jimeno, 2016).

We have three strategies to deal with this potential issue. First, the hazard analysis presented in Table 3 suggests that dryish counties are essentially a random sample from all dry counties, implying not only that their treatment is random, but also that there are no correlated missing regressors to contend with for this particular treatment. Second, as noted above, we include the interactions of all county-level time-invariant regressors used in the hazard analysis with time trends among the regressors. Since these regressors are known to influence the timing of becoming wet, their interaction with time-trends should pick up a substantial fraction of any county-level, time-varying factors that are correlated with the treatment. Third, since the bias we are concerned with results only from county-level, time-varying missing regressors, we also conduct our analysis on a very short panel from 1933 to 1936 and find essentially the same results (see Table 8). Since there is less time elapsed in this very short panel, there is less scope for changes in alcohol preferences at the county level in driving our results. Taken together,

these elements lend support to the idea that our estimated treatment effects can be interpreted as causal effects.

5. Estimating the Effects of Repeal on Infant Mortality

We now turn to estimating the effects of repeal on infant mortality. Our means of estimation is the fixed effects binomial regression detailed in section 4. Again, we are particularly interested in the effect of one county's decision to go wet on infant mortality within its own borders and on infant mortality in neighboring counties. Thus, we relate variation in infant mortality on a county-level basis to variation in dryish and wet status along with a large set of control variables detailed in section 3. Alternatively, what might matter is not only whether a neighboring county opts for wet status, but also how far away that county is. Appendix C explicitly incorporates the spatial distribution of counties into our empirical model. The results presented there are consistent with those presented in this section and are omitted here for expositional purposes.[11]

The leftmost columns of Table 4 show estimates including only wet variables as treatment regressors. Importantly, dryish counties are here pooled into the control group. Columns (1) through (3) explain infant mortality rates as a function of county and fixed effects (all columns), retail sales (columns 2 and 3), and other county controls[12] (column 3) along with a variable indicating whether a particular county switched to wet status in a given year (*wet in*

---

[11] There are further reasons addressed in Appendix C as to why we choose to not take any of the specifications incorporating the distance separating counties as our baseline. These relate to a relative lack of power when we incorporate the heterogeneity in distance; (2) distance measures invariably suffer from measurement error in the context of not knowing the distribution of population within the counties; and (3) it is unclear what functional form any distance measure should take as the decay rate for variables like dryish is unknown.
[12] These include all covariates in Table 1 except for New Deal spending, access to medical institutions, and the number of hospital beds. These are included in later specifications as a robustness exercise.

initial year or W_{ct*} as in section 4) and an additional variable indicating whether a particular county had previously switched to wet status (*wet in subsequent years* or W_{ct} as in section 4). Regardless of the specification used, the coefficients attached to both of these indicators are inconsistently signed, small in magnitude, and imprecisely estimated. As we shall see in a moment, this specification by ignoring potential cross-border policy externalities serves to mask the effects on infant mortality of both a county becoming wet and a county becoming dryish.

Accordingly, columns (4) through (6) of Table 4 separate bone dry and dryish counties. What is critical here is that the control group between the two sets of columns varies: for columns (1) through (3), the control group is all dry counties—that is, both bone dry and dryish counties—while for columns (4) through (6), the control group is only bone dry counties. Here, the coefficients attached to dryish status are consistently positive and statistically significant at conventional levels across all specifications in columns (4) through (6). We see that by successively adding proxies for county-level income and county-level controls interacted with linear trends the coefficient on dryish— both in the year of adoption and subsequent years—is diminished but not substantially so. Thus, we interpret our preferred estimate of roughly 0.03 in column (6) as representing the causal effect of one county's decision to go wet on its neighboring counties which corresponds to a 3% increase in baseline infant mortality for those counties affected.

In a related fashion, the decision to separate bone dry and dryish counties also has an important implication on the coefficient associated with wet status. In particular, its becomes larger and statistically significant by the time we reach our preferred specification in column (6). Previously, in columns (1) through (3) our control group—that is, all dry counties—were

contaminated by the presence of dryish counties. By separating out dryish counties, the coefficient on wet becomes larger relative to the new control group of bone dry counties.

Our estimates suggest that the policy externality associated with our estimates on dryish would amount to an additional 1.77 additional infant deaths per 1000 live births in dryish counties, representing an excess of 1,804 infant deaths across the six years from 1934 to 1939. Wet status is also associated with an increase in infant mortality of roughly 2%. Our estimates suggest that this would amount to an additional 1.18 additional infant deaths per 1000 live births in wet counties, representing an excess of 11,861 infant deaths across the six years from 1934 to 1939.[13] Cumulating across all affected counties and all years, these results imply 13,665 excess infant deaths that could be attributed to the repeal of federal prohibition in 1933.

These results highlight two of the main arguments of this paper. First, the distinction between bone dry and dryish counties turns out to be an important one. This finding suggests that cross-border policy externalities are likely important, both in contemporary and historical settings. Second, for whatever benefits the repeal of federal prohibition conferred in terms of consumer welfare, diminished expenditure on enforcement, and/or freedom of choice, it also came at the cost of increasing baseline infant mortality rates in both dryish and wet counties. Naturally, there were other associated costs which remain unexplored in this paper, but which should be added to any reckoning of repeal's legacy.

In what follows, we subject these results to a series of robustness checks. Critically, the results in Table 4 on the causal effects of dryish are remarkably robust across all specifications

---

[13] Table 2 reports an average infant mortality rate of 58.98 per 1000 live births in 1933. Our point estimate for dryish is roughly 0.03, translating into an additional 1.77 infant deaths per 1000 live births associated with dryish status. There were 1,019,239 live births in dryish counties from 1934 to 1939 which translates into an excess of 1,804 infant deaths in the same period. Our point estimate for wet is roughly 0.02, translating into an additional 1.18 infant deaths per 1000 live births associated with wet status. There were 10,051,925 live births in wet counties from 1934 to 1939 which translates into an excess of 11,861 infant deaths in the same period.

considered. The estimated causal effect of becoming wet are also relatively robust (though not quite as strongly as that of going dryish).

Are the estimated effects robust to further controls?

It is necessary to establish the robustness of our results under a wide range of specifications. To this end, Table 5 incorporates further controls. Specifically, column (1) replicates the results from column (6) in Table 4 for ease of comparison and further controls are added successively across columns. First, there may be concerns that the "treatment" status of dryish or wet may somehow be correlated with New Deal spending at the county-level which in turn may be correlated with infant mortality. With virtually unchanged coefficients on dryish and wet—both in initial and subsequent years—in column (2), this seems not to be the case. There may also be concern that the putative causal effect on dryish may reflect differential access to medical care—that is, dryish counties in particular simply had fewer resources (perhaps derived from revenues on alcohol sales) to check infant mortality. After including the number of medical institutions per capita and the number of hospital beds per capita, Column (3) suggests that this was not the case as, again, the coefficients are virtually unchanged.

We are also unable to control for county-level, time-varying unobservables through the use of county and year fixed effects. Of course, county-year fixed effects are infeasible, but county and year fixed effects along with state trends are not. Column (4) reports the results of this regression. For the most part, it mirrors the results for column (1). The only exception in this regard is the reduction in the size of the coefficient for *wet in subsequent years* and an increase in its standard error, yielding a statistically insignificant point estimate in this instance.

However, we do not put much interpretive weight on this result as the category of wet includes both counties which opt for wet status via county-level local elections and counties within states which opt for wet status via state-level legislation. By including state trends (or state-year fixed effects), we thereby eliminate any variation coming from the latter. Finally, we substitute state trends with state-year fixed effects in a regression with county and year fixed effects along with the whole host of control variables. Although the coefficient on *dryish in initial year* is slightly reduced in column (5), the coefficient on *dryish in subsequent years* are remarkably robust even after conditioning on a very large set of controls. Consequently, the results in column (5) are especially encouraging in that they control for any number of unobserved contemporaneous changes to government programs, legislative enactments, and local economic conditions that vary at the state-year level.

What if we only consider those counties which were at some point "treated"?

One potentially large, remaining concern may be that the omitted category—namely, bone dry—exhibits a different time trend than treated counties. This, of course, would violate the common trend assumption underlying our use of fixed effects—or equivalently, a difference-in-difference framework—and may indicate the influence of other unobserved county characteristics in driving our results. Accordingly, we can restrict estimation of our model to only those counties which experienced at least one change in treatment status over the years from 1933 to 1939, thereby, dropping counties which remain bone dry throughout and only exploiting the *timing* of the changes in prohibition status for the remaining counties.

Table 6 does precisely this replicating the various specifications of Table 5 but with a loss of 388 counties which remain bone dry throughout the period. Thus, column (1) does so for

our baseline specification. Column (2) incorporates per capita New Deal spending by county. Column (3) does the same but controls for differential access to medical institutions and hospital beds while columns (4) and (5) retain all of the controls but respectively add state trends and state-year fixed effects to the baseline controls of county and year fixed effects. Again, the wet variable wavers somewhat in its significance in columns (4) and (5). And as before, we do not put much interpretive weight on this result as the category of wet includes both counties which opt for wet status and counties within states which opt for wet status. By including state trends (or state-year fixed effects), we thereby eliminate any variation coming from wet states. Instead, we emphasize the remarkable resilience of the dryish variable across all five columns and of the wet variable across the first three columns with respect to size and significance of the coefficients.

What if we only consider neighboring county-pairs?

In an influential paper, Dube *et al*. (2010) propose the use of county-pairs which straddle state borders to assess the effects of minimum wage laws in the United States. This approach thereby exploits variation in the level of minimum wages induced by differential state legislation as in much of the earlier literature. However, it also chooses to use only neighboring counties, arguing that these are likely to be the most suitable control group given the similarity of local economic conditions across county borders.

In this spirit, we propose an alternative specification to estimate the effect of dryish in particular on infant mortality by limiting our sample only to dryish counties and their neighboring counties which were simultaneously bone dry in any given year. Thus, all wet counties in any given year are excluded and bone dry counties which are not neighbors with a

32

dryish county in any given year are also excluded. Implicitly, our rationale is much the same as in Dube *et al*. (2010): contiguous counties are likely to be the most suitable control group if a neighboring bone dry county is more similar to its dryish counterpart than to a randomly chosen bone dry county. However, a difference arises in that we compare dryish counties and neighboring (often interior) bone dry counties *within the same state* unlike Dube *et al*. (2010) which compare neighboring county-pairs that are located on opposite sides of a state border.

The advantage of this approach is that the control group—being geographically adjacent to dryish counties within the same state—is arguably more comparable to the treatment group. That is, the control group likely shares common, but unobserved county characteristics with the treatment group.[14] The downside of this approach is that we are likely to lose power as we are necessarily reducing our sample size by only considering dryish counties and their bone dry neighbors. On the contrary, using all US counties as in our preferred specification in column 6 of Table 4 is more nationally representative.

To begin, we identify all counties that ever became dryish during our sample period ($n = 700$). Then, for each such county, we match all the neighboring counties within the state, allowing for all possible combinations (that is, multiple matches for each dryish county as county lines do not necessarily line up perfectly). Third, we consider the prohibition status of each neighboring county and remove neighboring county-pairs for which we do not observe the combination of dryish/bone dry at any point in our sample period. The resulting sample size is 474 dryish counties which yield 985 distinct dryish/bone dry county-pairs.[15] Thus, on average,

---

[14] Our approach also allows for a further advantage: since our county-pairs are located in the same state, the constituent counties are commonly affected by any changes in state laws which are also unobserved.

[15] By using neighboring dryish/bone dry county-pairs, we are forced to drop a number of dryish counties from our sample because of the timing of their change in prohibition status. For example, if a neighbor of a dryish county becomes dryish in the same year, then this county-pair is dropped as it do not have the combination of dryish/bone

each dryish county has two neighboring counties which were bone dry for at least one year.

Finally, we remove those pair-year observations for which the bone dry/bone dry or dryish/bone

dry combination does not hold. This procedure yields 5,257 pair-years for a final observation

count of 10,514.[16]

Estimation takes place via MLE as before. However, in assessing infant mortality of

county c from county-pair p in time t, we include not only county and time fixed effects as

before, but also time-varying county-pair effects. The latter, then, implies that we are only using

variation in prohibition status within each contiguous county-pair to identify the effect of dryish

on infant mortality. Note that since our sample, in effect, "stacks" each county-pair, any

particular county will be in the sample as many times as it can be paired with a neighboring

county. The identifying assumption for this specification is that the prohibition status of counties

within a county-pair is uncorrelated with the differences in residual infant mortality in either

county. Naturally, this potential presence of counties in multiple pairs induces mechanical

correlation across county-pairs as discussed in Dube *et al*. (2010). To account for this, we cluster

standard errors at the state-year level, noting that there are 22 states and 108 state-years in our

sample (Cameron *et al*, 2011).

Column (3) in Table 7 (our baseline specification) shows that the estimate on dryish in

subsequent years is -0.030, which is very similar to the estimate of -0.027 in our preferred

specification from Table 4. This suggests that the control group in our preferred specification—

that is, all bone dry counties—is likely appropriate. However, we are able to achieve greater

dry needed in the estimation. Note, however, that such dryish counties are retained in our preferred specification as, there, we do not restrict our control group to be neighboring bone dry counties.

[16] This implies that each of the 985 contiguous county-pairs is observed 5.3 years on average, including the years in which both counties in the pair are bone dry. Alternatively, we can restrict our sample to only the county pairs that stayed dryish/bone dry for the entire sample period in order to have a balanced panel. The results are similar to those presented here and are omitted for expositional purposes (results available by request).

precision in our preferred specification in column 6 of Table 4 due to the larger sample size and

the need to estimate fewer parameters. This pattern is also seen in the other columns of Table 7

where the coefficient value for dryish is highly stable, but at best, statistically significant at the

10% level.

Are the estimated effects robust to other specifications?

Table 8 incorporates other specifications to further establish the robustness of the dryish

and wet effects. Specifically, column (1) replicates the results from column (6) in Table 4 for

ease of comparison and further checks on this specification are incorporated in turn. First,

consider the possibility that linear trends interacted with our county controls may be obscuring

important variation on a year-by-year basis which may be correlated with our dryish and wet

indicator variables. Therefore, it may be appropriate to incorporate more flexibility into our

specification and include our county controls interacted with year fixed effects. However, if

anything, the results of this exercise reported in column (2) suggest the contrary: the coefficients

on *dryish in subsequent years* and *wet in subsequent years* remain unaffected.

We also extend the sample up to 1941 in column (3).[17] Previously, we argued for a

terminal date of 1939, given that there is little variation in treatment status after that date. Here,

we consider a terminal date of 1941 to extend the sample but avoid any effects that World War II

and its associated mobilization effort might have had on infant mortality. The largest change

occurs for the coefficient on *wet in subsequent years* which becomes grossly statistically

insignificant while the results on dryish remain materially the same. Column (4) excludes 113

---

[17] Again, the choice of start date in 1933 is predicated in the main by the fact that mortality statistics for Texas, with its 254 counties, only begins in this year.

counties which border Canada or Mexico to account for the potential cross-board smuggling with materially similar results. We also consider the placebo effect of adding lead terms for the treatment variables in our preferred specification in column 6 of Table 4. That is, if a county becomes dryish in 1935, we assign a value of one for a new indicator variable for this county in 1934. For both dryish and wet status, no lead terms are individually or jointly statistically significant in column (5), suggesting that we are not picking up the residual effects of unobserved county characteristics in our preferred specification.

In column (6), we consider whether or not our results are visible in a shorter panel covering 1933 to 1936. The majority of switches to wet status occur over 1934 and 1935, so there is a fair amount of identifying variation here. We see that our main results from Table 4 are roughly consistent with the results from the shorter panel as we find positive effects on the order of 2−3% for our wet and dryish treatments, but with somewhat larger standard errors. In a shorter panel, there is necessarily less time for county-level, time-varying correlated missing regressors to induce bias. The implication here is that the bias due to such correlates is likely not very large, given the similarity of results across Tables 4 and 8. Thus, the estimates in column (6), while imprecise, reassure us that the inclusion of county fixed-effects does a reasonable job of accounting for the fact that preferences for alcohol and the decision to become wet ought to be correlated and that such preferences are relatively stable over time.

What about wet states as opposed to wet counties?

Up to this point, there have been two consistently documented results. The plausibly exogenous status of dryish is systematically associated with higher infant mortality rates. At the same time, we find a county's status as wet is also systematically associated with higher infant

mortality rates, provided that state-trends and state-year fixed effects are not included. One curious feature, though, of these results in combination is that the coefficient for *dryish in subsequent years* is always greater than that for *wet in subsequent years*. While our analytical framework from section 2.2 gives some reason to think that this is a plausible set of results, Table 9 explores this issue in greater depth.

Here, we make the distinction in between those counties which went wet through local option ("wet county") and those which went wet through state legislation ("wet state"). The reason for doing so is that the latter changes in prohibition status are arguably more exogenous than the former from the perspective of individual counties. That is, a county's inhabitants could have strong preferences for remaining dry but find themselves residing in a state with strong preferences for becoming wet. Thus, such wet counties may be rough analogs to their dryish counterparts. Column (1) makes this distinction for our baseline specification. Column (2) incorporates per capita New Deal spending by county. Column (3) does the same but controls for differential access to medical institutions and hospital beds.

Perhaps unsurprisingly, very little changes as it relates to the results on dryish. However, some interesting results emerge for wet states. In particular, the coefficients for *wet state in subsequent years* are, for the most part, statistically significant and virtually indistinguishable in magnitude from *dryish in subsequent years*. This is especially true as we move from column (1) to column (3) as the difference between the two coefficients becomes truly negligible. We treat this as a piece of corroborating evidence for our estimates of the causal effect on infant mortality of dry counties becoming dryish and why they may be larger than those for counties becoming wet in our preferred specification.

6. Conclusion

In considering the effects of the repeal of federal prohibition, we find robust evidence that relaxing restrictions on alcohol sales lead to increases in infant mortality. Critical in establishing this result is recognizing that it is not only an individual county's choice of prohibition status which matters but also the prohibition status of its neighbors. Thus, our strongest set of results—both in the estimated magnitude of the effect and in the number of specifications for which it holds—relates to dry counties being "treated" with wet neighbors. Clearly, this type of policy externality is important not only in the context of assessing the repeal of federal prohibition but also in the context of current policy debates related to states potentially legalizing other illicit substances.

This paper also documents that these developments occurred in an environment when the general trajectory of infant mortality rates was distinctly downward. From 1934 to 1939, the infant mortality rate for the US dropped from 60.0 per thousand live births to 50.2 per thousand in 1939, or by 16.33%. We estimate that dryish status was associated with a 3% increase in infant mortality rates, a number which we have argued before likely represents an unbiased estimate of the causal effects of a dry county being treated with a wet neighbor. We also estimate that wet status was associated with a 2% increase in infant mortality rates. Thus, the repeal of federal prohibition can be thought of having reversed the generalized decline in infant mortality rates in this period by 12.25 to 18.37% in treated counties.[18] Again, cumulating across all affected counties and all years, these results imply a minimum of 13,665 excess infant deaths that could be attributed to the repeal of federal prohibition in 1933.

---

[18] These figures are simply calculated as the ratio of the 2 or 3% increase in infant mortality rates for treated counties to the 16.33% decline in the infant mortality rate for the entire US over the same period.

We have been relatively silent on mechanisms, instead offering a preferred interpretation of the data in the form of potential maternal alcohol consumption. And while there is an established medical literature which suggests a link from maternal alcohol consumption to infant mortality via compromised immune systems and low birth-weights, we have very little by way of corroborating evidence in support of this hypothesis. Thus, other linkages in between the availability of alcohol and infant death remain as possibilities and as an area for future work.

Further avenues for future work come in assessing the effects of repeal on other contemporaneous outcomes, such as adult morbidity and mortality, violent crime, and worker productivity. Similar work could also exploit the variation in prohibition laws at the county level which predated federal prohibition in 1920 and which has been neglected in the literature. More ambitiously, we hope to explore the long-run effects of prohibition by considering how changes in potential maternal alcohol consumption induced by prohibition laws affected children born in these periods throughout their lives. Thus, we will correlate the ample geographic and temporal heterogeneity in restrictions on alcohol, both before and after federal prohibition, with long-term outcomes such as educational attainment, occupational status, and wages. Taken together, such work will—at last—allow a final tab for prohibition in all of its forms to be drawn.

References

Acemoglu, D., D.H. Autor, and D. Lyle (2004), "Women, War, and Wages: The Effect of Female Labor Supply on the Wage Structure at Midcentury." *Journal of Political Economy* 112(3), 497-551.

Alsan, M. and C. Goldin (2015), "Watersheds in Infant Mortality: The Role of Effective Water and Sewerage Infrastructure, 1880 to 1915." *NBER Working Paper 21263*.

Anand, S. and T. Bärnighausen (2004), "Human Resources and Health Outcomes: Cross-country Econometric Study." *The Lancet* 364(9445), 1603-1609.

Baird, S., J. Friedman, and N. Schady (2011), "Aggregate Income Shocks and Infant Mortality in the Developing World." *Review of Economics and Statistics* 93(3), 847-856.

Barreca, A. and M. Page (2015), "A Pint for a Pound? Reevaluating the Relationship Between Minimum Drinking Age Laws and Birth Outcomes." *Health Economics* 24(4), 400-418.

Blocker, J.S. (2006), "Did Prohibition Really Work? Alcohol Prohibition as a Public Health Innovation." *American Journal of Public Health* 96(2), 233-243.

Bodenhorn, H. (2016), "Blind Tigers and Red-Tape Cocktails: Liquor Control and Homicide in Late-Nineteenth Century South Carolina." *NBER Working Paper 22980*.

Cameron, A.C., J.B. Gelbach, and D.L. Miller (2011), "Robust Inference With Multiway Clustering." *Journal of Human Resources* 29(2), 238-249.

Carpenter, C. and C. Dobkin (2009), "The Effect of Alcohol Consumption on Mortality: Regression Discontinuity Evidence from the Minimum Drinking Age." *American Economic Journal: Applied Economics* 1(1), 164-182.

Clay, K., J. Lewis, and E. Severnini (2016), "Canary in a Coal Mine: Infant Mortality, Property Values, and Tradeoffs Associated with Mid-20th Century Air Pollution." *NBER Working Paper 22155*.

Cook, P. (2007), *Paying the Tab: The Costs and Benefits of Alcohol Control*. Princeton: Princeton University Press.

Culver, D.C. and J.E. Thomas (1940), *State Liquor Control Administration*. Bureau of Public Administration, University of California-Berkeley.

Cutler, D. and G. Miller (2005), "The Role of Public Health Improvements in Health Advances: The Twentieth-Century United States." *Demography* 42(1), 1-22.

Dills, A.K. and J.A. Miron (2004), "Alcohol Prohibition and Cirrhosis." *American Law and Economics Review* 6(2), 285-318.

Distilled Spirits Institute (1935), *Summary of State Liquor Control Laws and Regulations Relating to Distilled Spirits*. Washington, D.C.: Distilled Spirits Institute.

Distilled Spirits Institute (1941), *Summary of State Liquor Control Laws and Regulations Relating to Distilled Spirits*. Washington, D.C.: Distilled Spirits Institute.

Doyle, J.J. and K. Samphantharak (2008), "$2.00 Gas! Studying the Effects of a Gas Tax Moratorium." *Journal of Public Economics* 92(4), 869-884.

Dube, A., T.W. Lester, and M. Reich (2010), "Minimum Wage Effects Across State Borders." *Review of Economics and Statistics* 92(4), 945-964.

Evans, M., E. Helland, J. Klick, and A. Patel (2016), "The Developmental Effect of State Alcohol Prohibitions at the Turn of the 20th Century." *Economic Inquiry* 54(2), 762-777.

Fishback, P.V., M. Haines, and S. Kantor (2001), "The Impact of the New Deal on Black and White Infant Mortality in the South." *Explorations in Economic History* 38(1), 93-122.

Fishback, P.V., M. Haines, and S. Kantor (2007), "Births, Deaths, and New Deal Relief During the Great Depression." *Review of Economics and Statistics* 89(1), 1-14.

Fishback, P.V., W. Troesken, T. Kollmann, M. Haines, P. Rhode, and M. Thomasson (2011), "Information and the Impact of Climate and Weather on Mortality Rates During the Great Depression." In *The Economics of Climate Change* (Ed.s G. Libecap and R. Steckel). Chicago: University of Chicago Press, 131-168.

García-Jimeno, C. (2016), "The Political Economy of Moral Conflict: An Empirical Study of Learning and Law Enforcement under Prohibition." *Econometrica* 84(2), 511-570.

Hahn, J. and W. Newey (2004), "Jackknife and Analytical Bias Reduction for Nonlinear Panel Models." *Econometrica* 72(4), 1295-1319.

Harding, M., E. Leibtag, and M.F. Lovenheim (2012), "The Heterogeneous Geographic and Socioeconomic Incidence of Cigarette Taxes: Evidence from Nielsen Homescan Data." *American Economic Journal: Economic Policy* 4(4), 169-198.

Harrison, L.V. (1938), *The Local Option Fallacy*. Washington, D.C.: Distilled Spirits Institute.

Horrace, W.C. and R.L. Oaxaca (2006), "Results on the Bias and Inconsistency of Ordinary Least Squares for the Linear Probability Model." *Economics Letters* 90(3), 321-327.

Hoynes, H.W. and D.W. Schanzenbach (2012), "Work Incentives and the Food Stamp Program." *Journal of Public Economic* 96(1), 151-162.

Johansson, P., T. Pekkarinen, and J. Verho (2014), "Cross-border Health and Productivity Effects of Alcohol Policies." *Journal of Health Economics* 36(1), 125-136.

Kyvig, D.E. (2000), *Repealing National Prohibition*. Ashland: Kent State University Press.

Lovenheim, M.F. (2008), "How Far to the Border? The Extent and Impact of Cross-Border Casual Cigarette Smuggling." *National Tax Journal* 61(1), 7-33.

Lovenheim, M.F. and J. Slemrod (2010), "The Fatal Toll of Driving to Drink: The Effect of Minimum Legal Drinking Age Evasion of Traffic Fatalities." *Journal of Health Economics* 29(1), 62-77.

Machado, M.P. (2004), "A Consistent Estimator for the Binomial Distribution in the Presence of 'Incidental Parameters'." *Journal of Econometrics* 119(1), 73-98.

Mills, J.L., B.I. Graubard, E.E. Harley, G.G. Rhoads, and H.W. Berendes (1984), "Maternal Alcohol Consumption and Birth Weight. How Much Drinking during Pregnancy is Safe?" *Journal of the American Medical Association* 252(14), 1875-1879.

Miron, J.A. and J. Zwiebel (1991), "Alcohol Consumption during Prohibition." *American Economic Review, Papers & Proceedings* 81(2), 242-247.

Nilsson, P.J. (2016), "Alcohol Policy, Prenatal Conditions, and Long-term Economic Outcomes." *Journal of Political Economy*, forthcoming.

Okrent, D. (2010), *Last Call: The Rise and Fall of Prohibition*. New York: Scribner.

Olegård, R., K.G. Sabel, M. Aronsson, B. Sandin, P.R. Johansson, C. Carlsson, M. Kyllerman, K. Iversen, and A. Hrbek (1979), "Effects on the Child of Alcohol Abuse during Pregnancy: Retrospective and Prospective Studies." *Acta Paediatrica* 68(S275), 112-121.

Stepner, M. (2016), "binscatter: A stata program to generate binned scatterplots" Available at https://michaelstepner.com/binscatter/ [accessed on January 17, 2017].

Strandberg-Larsen, K., M. Grønbæk, A. Andersen, P. Andersen, and J. Olsen (2009). "Alcohol Drinking Pattern During Pregnancy and Risk of Infant Mortality." *Epidemiology* 20(6), 884-891.

Strumpf, K.S. and F. Oberholzer-Gee (2002), "Endogenous Policy Decentralization: Testing the Central Tenet of Economic Federalism." *Journal of Political Economy* 110(1), 1-36.

Tan, C.H., C.H. Denny, N.E. Cheal, J.E. Sniezek, and D. Kanny (2015), "Alcohol Use and Binge Drinking among Women of Childbearing Age–United States, 2011-2013." *Morbidity and Mortality Weekly Report* 64(37), 1042-1046.

Thomas, J.E. and D.C. Culver (1940), "Protection of Dry Areas." *Law and Contemporary Problems* 7(4), 696-708.

Figure 1a: A County with High Willingness-to-pay (WTP) for Alcohol
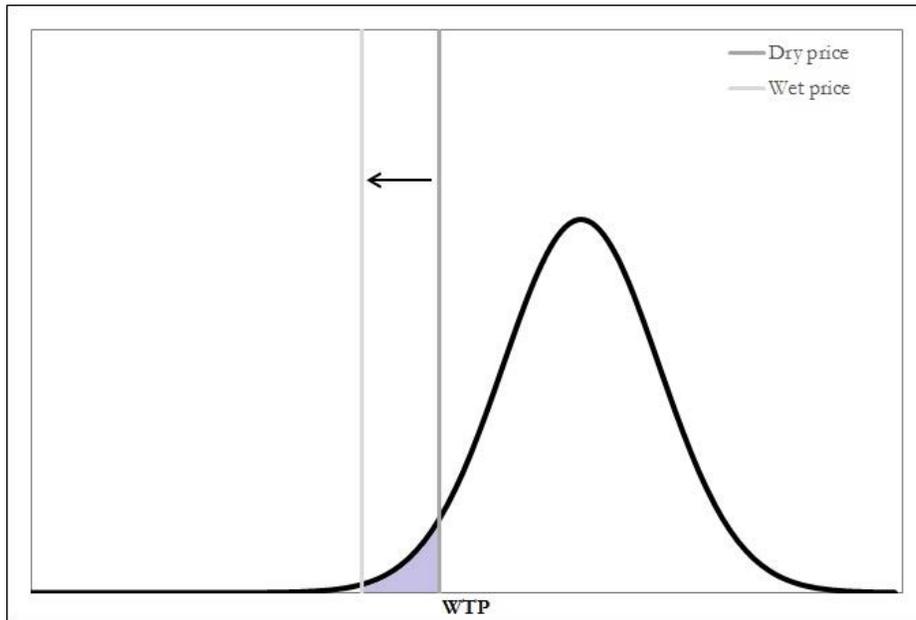
Figure 1b: A County with Low Willingness-to-pay (WTP) for Alcohol
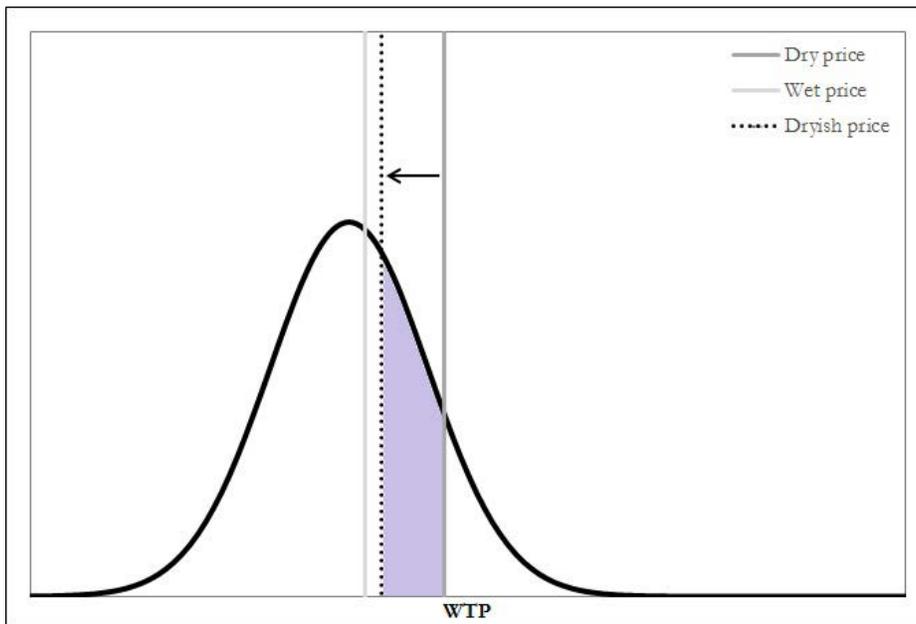
Notes: We assume that each county can be characterized by its distribution of individuals' willingness-to-pay (WTP) for alcohol and that this distribution differs by county. To simplify, let the distribution of WTP be symmetric and unimodal so that the median voter is at the top (mode) of the density function. For individuals, assume that consuming alcohol at a dangerous level is a binary decision. They do so if their willingness-to-pay for alcohol exceeds its price. The effective price of purchasing and consuming alcohol while the county is variously dry, dryish, or wet are respectively denoted by "dry price", "dryish price", and "wet price".

Figure 2: Infant Mortality Rates by Prohibition Status, 1934–1939 (deaths per 1000 births)
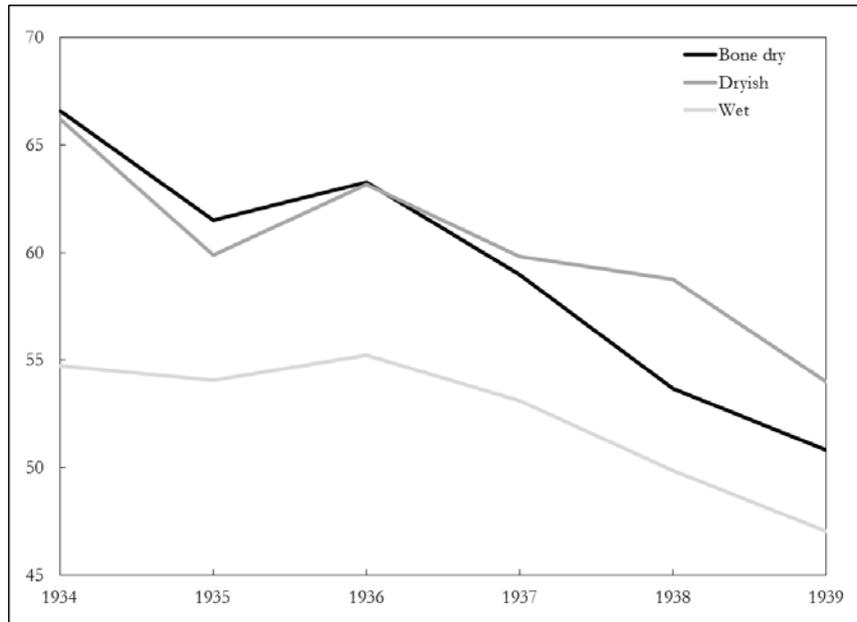
Notes: The infant mortality rate is the number of infant deaths within a year from live birth per 1000 births. The number of births for each county are used as weights. *Bone dry* are dry counties surrounded by other dry counties. *Dryish* counties are dry counties with at least one wet neighbor. *Wet* counties are counties which allow for alcohol sales within their borders. There are 3,043 counties in our sample.

Figure 3: Proportion of US Counties by Prohibition Status, 1930–1942
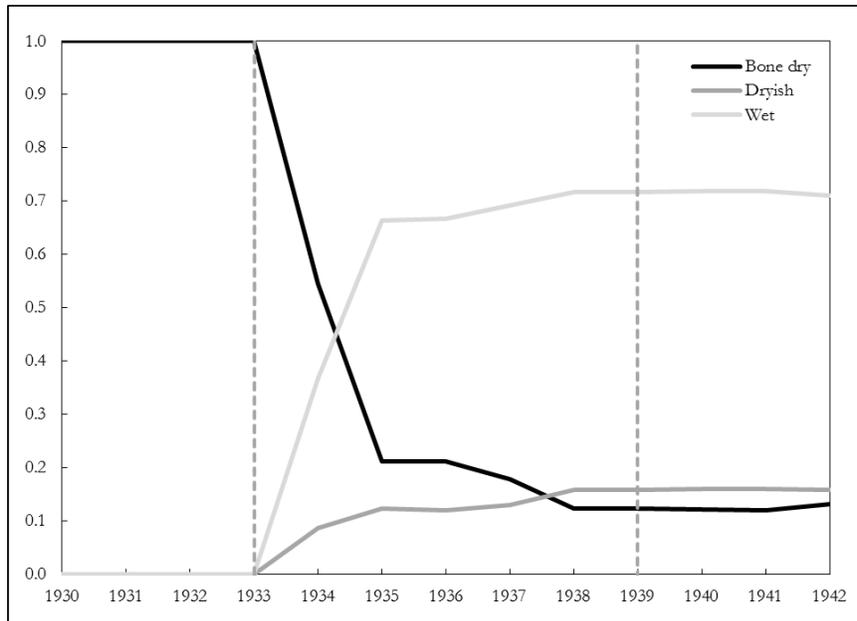
Notes: Bone dry counties are dry counties surrounded by other dry counties. Dryish counties are dry counties with at least one wet neighbor. The figure treats every county as bone dry in 1933. The two vertical dashed lines correspond to the beginning (1933) and end (1939) of our sample period. Dryish counties are dry themselves but have at least one wet neighbor. There are 3,043 counties in our sample.

Figure 4: Spatial Distribution of Counties by Prohibition Status

1933

1934

1935

1936

1937

1938

1939
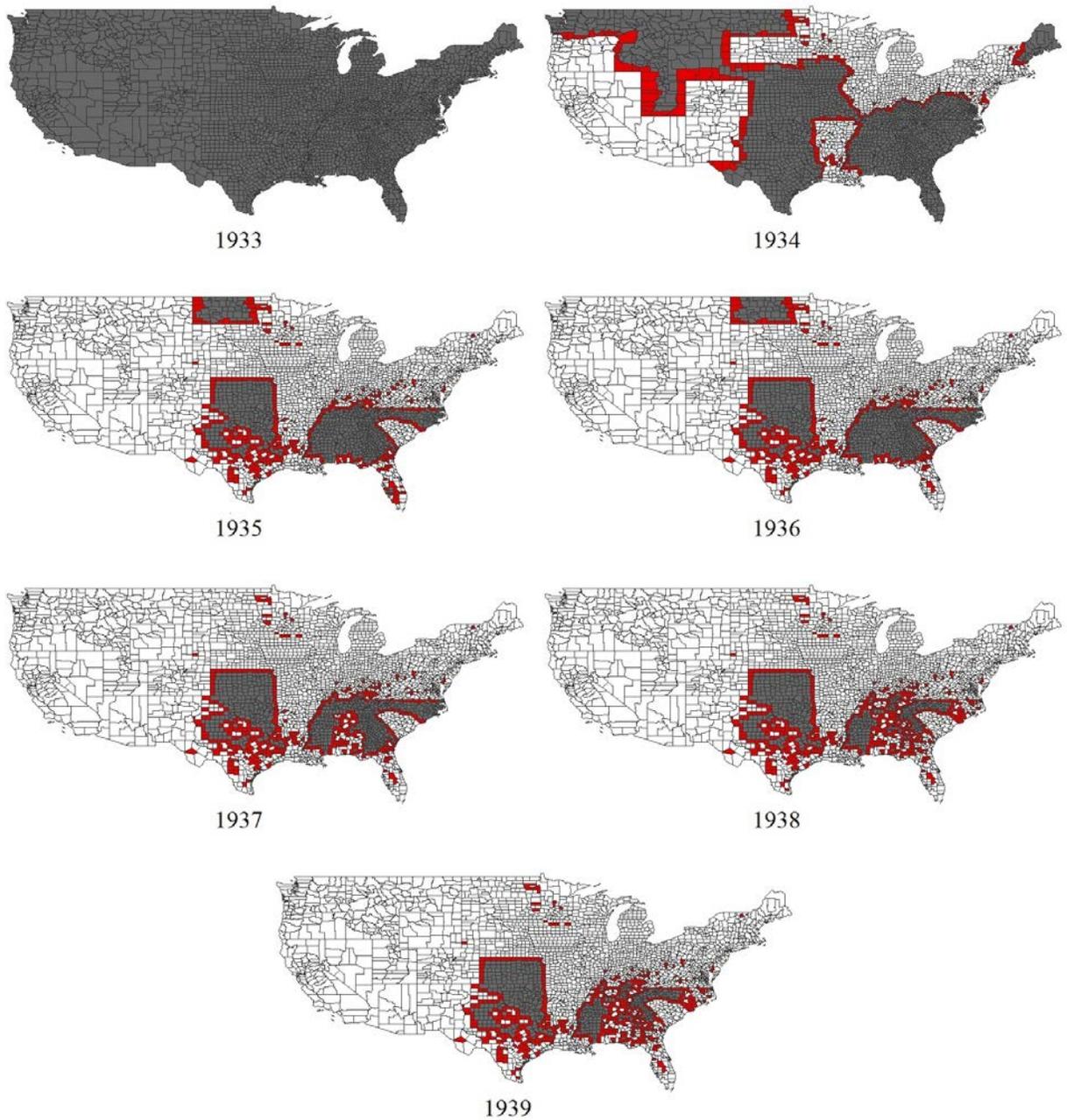
Notes: The counties in dark gray, red, and white correspond to bone dry, dryish, and wet counties, respectively. Bone dry counties are dry counties surrounded by other dry counties. Dryish counties are dry counties with at least one wet neighbor.

Figure 5: Pre-trends in Infant Mortality Rates, 1928–1933 (deaths per 1000 births)
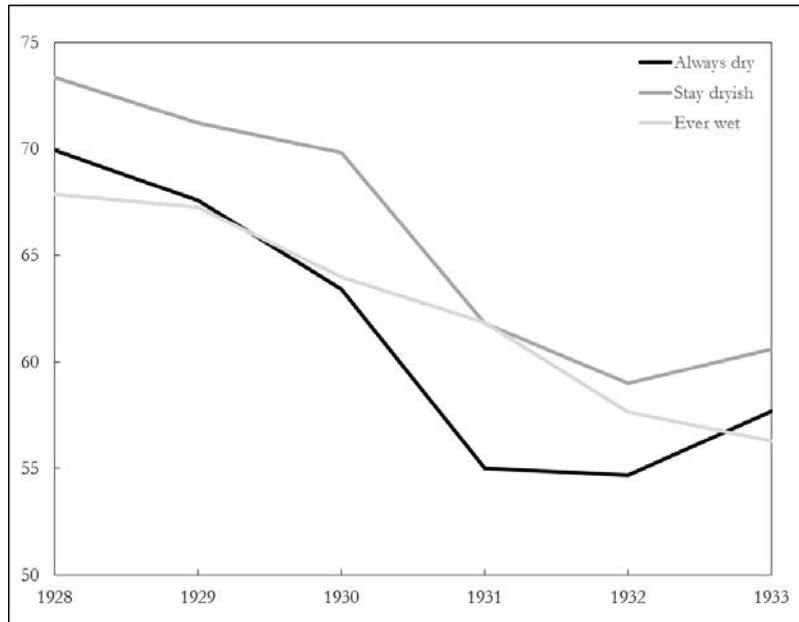
Notes: The infant mortality rate is the number of infant deaths within a year from live birth per 1000 births. The number of births for each county are used as weights. *Always dry* are counties that stayed bone dry during our sample period from 1933 to 1939. *Stay dryish* are counties that became dryish at any time from 1933 to 1939 and stay dryish until 1939. *Ever wet* are counties that became wet at any time from 1933 to 1939. As such, counties that became dryish first but eventually became wet are included in *ever wet* and not in *stay dryish*. Thus, the three categories are mutually exclusive. The sample underlying this figure is limited to the 2,670 counties which can be traced back to 1928. The choice of 1928 is predicated by the fact that sample size shrinks dramatically as we add earlier years.

Table 1: Variable Definitions and Data Sources

Variable name	Definition	Source
Retail sales	Retail sales per capita, linearly interpolated between 1933, 1935, and 1939 (*time varying*)	Fishback *et al.* (2011)
% black	Number of blacks divided by total population in 1930	1930 Census - State and County I
% urban	Number of urban residents divided by total population in 1930	1930 Census - State and County I
% foreign	Number of foreign born divided by total population in 1930	1930 Census - State and County I
% female	Number of females divided by total population in 1930	1930 Census - State and County I
% Protestant	Number of Protestants in 1926 divided by total population in 1930	1926 Census of Religious Bodies
% childbearing age	Number of females aged 15-44 divided by total female population in 1930	1930 Census - State and County I
Unemployment rate	Number of unemployed divided by population aged 15-64 in 1930	1930 Census - State and County I
Log(population)	Log of total population in 1930	1930 Census - State and County I
New Deal spending	Cumulative New Deal spending from March 1933 through June 1939 divided by total population in 1930	Fishback *et al.* (2011)
Access to medical institutions	Number of medical institutions divided by total population in thousands (*time varying*)	Fishback *et al.* (2011)
Hospital beds	Hospital beds per 1000 women aged 15-44 (*time varying*)	Fishback *et al.* (2011)

Sources: Fishback, P.V., W. Troesken, T. Kollmann, M. Haines, P. Rhode, and M. Thomasson (2011), "Information and the Impact of Climate and Weather on Mortality Rates During the Great Depression." In *The Economics of Climate Change* (Ed.s G. Libecap and R. Steckel). Chicago: University of Chicago Press, 131-168; Gardner, J. and W. Cohen (1992), "Demographic Characteristics of the Population of the United States, 1930-1950: County-Level." Ann Arbor: Inter-university Consortium for Political and Social Research, http://doi.org/10.3886/ICPSR00020.v1; US Department of Commerce, Bureau of the Census (1980), *Censuses of Religious Bodies, 1906-1936*. Ann Arbor: Inter-university Consortium for Political and Social Research, http://doi.org/10.3886/ICPSR00008.v1

Table 2: Baseline Sample County Characteristics by Treatment Group

	All	Always dry	Stay dryish	Ever wet
		Mean by treatment group		
	(1)	(2)	(3)	(4)
Infant mortality rate in 1933	57.93	58.00	61.17	57.55
	[19.48]	[17.82]	[20.76]	[19.46]
Log (infant mortality rate) in 1933	4.01	4.01	4.06	4.01
	[0.31]	[0.31]	[0.34]	[0.30]
Retail sales in 1933	483.93	290.56	235.58	532.63
	[228.23]	[165.14]	[119.87]	[213.67]
% black	10.54	16.65	22.34	8.54
	[16.19]	[20.87]	[22.66]	[13.83]
% urban	49.15	22.11	13.03	56.12
	[36.96]	[25.35]	[18.90]	[35.82]
% immigrant	9.62	0.78	0.68	11.56
	[10.19]	[1.39]	[2.18]	[10.24]
% childbearing age	45.00	40.84	38.37	46.20
	[11.95]	[16.21]	[16.93]	[10.27]
% Protestant in 1926	28.48	37.33	40.64	26.15
	[13.25]	[11.79]	[12.58]	[12.28]
% female	46.52	43.01	41.69	47.44
	[11.66]	[16.74]	[18.22]	[9.67]
Log (population)	10.91	8.95	8.54	11.39
	[3.19]	[3.55]	[3.77]	[2.86]
Unemployment rate	3.07	1.38	1.03	3.47
	[2.11]	[1.32]	[1.15]	[2.04]
New Deal spending	156.03	111.19	94.26	167.81
	[147.20]	[78.24]	[62.40]	[156.67]
Access to medical institutions	57.17	45.84	35.33	60.87
	[50.45]	[45.26]	[41.84]	[51.06]
Hospital beds	14.16	7.37	6.32	15.77
	[13.73]	[9.30]	[19.30]	[12.78]
Observations	3,043	388	475	2,180

Notes: Column (1) reports means for the full sample while columns (2)–(4) report means by each treatment group with standard deviations in brackets. The number of births in each county are used as weights. Always dry are counties that stayed bone dry during our sample period from 1933 to 1939. Stay dryish are counties that became dryish at any time from 1933 to 1939 and stay dryish until 1939. Ever wet are counties that became wet at any time from 1933 to 1939. As such, counties that became dryish first but eventually became wet are included in ever wet and not in stay dryish. Thus, the three categories are mutually exclusive. Unless mentioned, each variable in the table comes from 1930.

Table 3: Determinants of Prohibition Status by Predetermined County Characteristics

	Probability (wet)		Probability (dryish)	
	(1)	(2)	(3)	(4)
Retail sales in 1933	0.001***	0.001***	-0.001***	-0.001***
	(0.000)	(0.000)	(0.000)	(0.000)
% black	-0.004**	-0.003**	0.002	0.002
	(0.001)	(0.001)	(0.002)	(0.002)
% urban	-0.004***	-0.004***	0.002	0.001
	(0.001)	(0.001)	(0.003)	(0.003)
% immigrant	0.027***	0.026***	-0.018	-0.016
	(0.003)	(0.003)	(0.012)	(0.013)
% childbearing age	-0.064***	-0.068***	0.026**	0.029**
	(0.005)	(0.005)	(0.013)	(0.014)
% Protestant in 1926	-0.013***	-0.013***	-0.001	-0.001
	(0.001)	(0.001)	(0.003)	(0.003)
% female	0.036***	0.036***	-0.018	-0.019
	(0.006)	(0.006)	(0.014)	(0.015)
Log (population)	0.152***	0.167***	-0.077	-0.087
	(0.022)	(0.023)	(0.064)	(0.064)
Unemployment rate	0.110***	0.106***	-0.048	-0.048
	(0.012)	(0.012)	(0.043)	(0.043)
New Deal spending		0.166*		-0.249
		(0.091)		(0.378)
Access to medical institutions		0.530**		-0.335
		(0.212)		(0.734)
Hospital beds		-0.000		0.001
		(0.002)		(0.003)
Log likelihood	-16,648	-16,646	-5,271	-5,271
Observations	3,043	3,043	3,043	3,043

Notes: Estimates from a Cox hazard model are reported with robust standard errors in parentheses. The baseline year is 1933 when Federal Prohibition is repealed. Out of 3,043 counties, 388 counties stayed dry from 1933 to 1939 while 700 and 2,180 counties respectively became dryish and wet at some point in between 1933 to 1939. Out of 700 counties that were ever dryish, 225 counties in our sample became dryish first but eventually became wet. All explanatory variables are predetermined county characteristics from before 1933. Unless mentioned, each variable in the table comes from 1930. Significance levels: *** p<0.01, ** p<0.05, * p<0.1

Table 4: The Effect of Repeal on Infant Mortality Rates

	Without dryish			With dryish		
	(1)	(2)	(3)	(4)	(5)	(6)
Dryish in initial year				0.038***	0.034***	0.031***
				(0.012)	(0.012)	(0.012)
Dryish in subsequent years				0.044***	0.038***	0.027**
				(0.011)	(0.011)	(0.011)
Wet in initial year	0.003	0.004	-0.002	0.013	0.013	0.005
	(0.008)	(0.008)	(0.007)	(0.009)	(0.009)	(0.008)
Wet in subsequent years	-0.010	-0.001	0.006	0.010	0.015	0.018**
	(0.008)	(0.009)	(0.008)	(0.010)	(0.010)	(0.009)
Log likelihood	-62,765	-62,737	-62,597	-62,749	-62,725	-62,589
N	21,291	21,291	21,291	21,291	21,291	21,291
N of county	3,043	3,043	3,043	3,043	3,043	3,043
County FEs	X	X	X	X	X	X
Year FEs	X	X	X	X	X	X
Retail sales		X	X		X	X
Other county controls			X			X

Notes: Estimates from fixed effect maximum likelihood estimation (MLE) are reported with robust standard errors in parentheses. The sample size is 21,291 (3,043 counties for each year from 1933 to 1939 excluding 10 observations with no births in the year). Columns (1)–(3) do not distinguish dryish from dry while columns (4)–(6) separate dryish counties from dry counties. Critically, the control group between the two sets of columns varies: for columns (1)–(3), the control group is all dry counties while for columns (4)–(6), the control group is only bone dry counties. *Retail sales* are county-level retail sales per capita which is used as a proxy for income. Other county controls are the variables reported in columns (1) and (3) of Table 3 interacted with a linear trend, except for *retail sales* which is time-varying. Significance levels: *** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table 5: The Effect of Repeal on Infant Mortality Rates, Additional Controls

	(1)	(2)	(3)	(4)	(5)
Dryish in initial year	0.031***	0.031***	0.031***	0.030**	0.021*
	(0.012)	(0.012)	(0.012)	(0.012)	(0.012)
Dryish in subsequent years	0.027**	0.026**	0.026**	0.024**	0.024*
	(0.011)	(0.011)	(0.011)	(0.011)	(0.013)
Wet in initial year	0.005	0.005	0.005	0.003	0.013
	(0.008)	(0.008)	(0.008)	(0.008)	(0.016)
Wet in subsequent years	0.018**	0.018**	0.019**	0.013	0.001
	(0.009)	(0.009)	(0.009)	(0.010)	(0.018)
Log likelihood	-62,589	-62,582	-62,581	-62,446	-61,860
N	21,291	21,291	21,291	21,291	21,291
N of county	3,043	3,043	3,043	3,043	3,043
County FEs	X	X	X	X	X
Year FEs	X	X	X	X	X
Retail sales	X	X	X	X	X
Other county controls	X	X	X	X	X
New Deal spending		X	X	X	X
Medical institutions			X	X	X
Hospital beds			X	X	X
State-trends				X	
State-year FEs					X

Notes: Estimates from fixed effect maximum likelihood estimation (MLE) are reported with robust standard errors in parentheses. The sample size is 21,291 (3,043 counties for each year from 1933 to 1939 excluding 10 observations with no births in the year). Column (1) replicates our baseline estimates from Column (6) of Table 4. *Retail sales* are county-level retail sales per capita which is used as a proxy for income. Other county controls are the variables reported in columns (1) and (3) of Table 3 interacted with a linear trend, except for *retail sales* which is time-varying. *New Deal spending* is the cumulative amount of county-level New Deal spending per capita interacted with a linear trend. *Medical institutions* is the number of medical institutions per 1000 people in a county while *hospital beds* is the number of hospital beds per 1000 women aged 15–44 in a county (both of which are time-varying). Significance levels: *** p<0.01, ** p<0.05, * p<0.1

Table 6: The Effect of Repeal on Infant Mortality Rates, Only Treated in Sample

	(1)	(2)	(3)	(4)	(5)
Dryish in initial year	0.034***	0.034***	0.034***	0.032***	0.031**
	(0.012)	(0.012)	(0.012)	(0.012)	(0.014)
Dryish in subsequent years	0.031**	0.030**	0.030**	0.030**	0.038*
	(0.012)	(0.012)	(0.012)	(0.013)	(0.020)
Wet in initial year	0.008	0.008	0.008	0.005	0.024
	(0.008)	(0.008)	(0.008)	(0.008)	(0.019)
Wet in subsequent years	0.021**	0.021**	0.021**	0.012	0.009
	(0.011)	(0.010)	(0.010)	(0.011)	(0.026)
Log likelihood	-54,677	-54,669	-54,668	-54,532	-54,012
N	18,575	18,575	18,575	18,575	18,575
N of county	2,655	2,655	2,655	2,655	2,655
County FEs	X	X	X	X	X
Year FEs	X	X	X	X	X
Retail sales	X	X	X	X	X
Other county controls	X	X	X	X	X
New Deal spending		X	X	X	X
Medical institutions			X	X	X
Hospital beds			X	X	X
State-trends				X	
State-year FEs					X

Notes: Estimates from fixed effect maximum likelihood estimation (MLE) are reported with robust standard errors in parentheses. We exclude those counties which never experienced a change in treatment status; that is, we exclude counties which were bone dry in every year from 1933 to 1939 (388 counties). The resulting sample size is 18,575 (again, excluding 10 observations with no reported births in the year). See Table 5 for the definition of each control variable. Significance levels: *** p<0.01, ** p<0.05, * p<0.1

Table 7: The Effect of Repeal on Infant Mortality Rates, Neighboring County-Pair Sample

	(1)	(2)	(3)	(4)	(5)
Dryish in initial year	-0.001	0.001	0.001	0.001	0.001
	(0.016)	(0.016)	(0.016)	(0.016)	(0.016)
Dryish in subsequent years	0.027	0.028*	0.030*	0.029*	0.029
	(0.017)	(0.017)	(0.018)	(0.018)	(0.018)
Log likelihood	-27,028	-27,023	-27,014	-27,013	-27,012
N	10,514	10,514	10,514	10,514	10,514
N of dryish counties	474	474	474	474	474
N of county pairs	985	985	985	985	985
N of county pair-years	5,257	5,257	5,257	5,257	5,257
County FEs	X	X	X	X	X
Year FEs	X	X	X	X	X
County pair-year FEs	X	X	X	X	X
Retail sales		X	X	X	X
Other county controls			X	X	X
New Deal spending				X	X
Medical institutions					X
Hospital beds					X

Notes: Estimates from fixed effect maximum likelihood estimation (MLE) are reported with standard errors clustered at state-year level in parentheses. We retain only those counties which were dryish in any given year and their neighboring counties which were simultaneously bone dry; that is, all wet counties in any given year are excluded and bone dry counties which are not neighbors with a dryish county in any given year are also excluded. The resulting sample size is 10,514 with 474 dryish counties which yield 985 dryish-bone dry county pairs. See Table 5 for the definition of each control variable. Significance levels: *** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table 8: The Effect of Repeal on Infant Mortality Rates, Additional Specifications

	(1)	(2)	(3)	(4)	(5)	(6)
	Baseline (column 6 of Table 4)	(1) with controls interacted with year FEs	(1) with extended sample to 1941	(1) w/o counties bordering Canada or Mexico	(1) with lead terms	(1) with limited sample to 1936
Dryish in initial year	0.031***	0.032***	0.033**	0.030**	0.040***	0.027**
	(0.012)	(0.012)	(0.012)	(0.012)	(0.013)	(0.014)
Dryish in subsequent years	0.027**	0.029**	0.031***	0.028**	0.036***	0.022
	(0.011)	(0.011)	(0.010)	(0.011)	(0.013)	(0.015)
Wet in initial year	0.005	0.016**	0.001	0.003	0.009	0.003
	(0.008)	(0.008)	(0.008)	(0.007)	(0.010)	(0.008)
Wet in subsequent years	0.018**	0.019**	0.009	0.019**	0.022*	0.019*
	(0.009)	(0.009)	(0.008)	(0.008)	(0.011)	(0.011)
Dryish one year before (t-1)					0.020	
					(0.014)	
Wet one year before (t-1)					0.002	
					(0.009)	
Log likelihood	-62,589	-62,389	-81,079	-60,031	-62,587	-34,939
N	21,291	21,291	27,384	20,500	21,291	12,172
County FEs	X	X	X	X	X	X
Year FEs	X	X	X	X	X	X
Retail sales	X	X	X	X	X	X
Other county controls with linear trends	X		X	X	X	X
Other county controls with year FEs		X				

Notes: Estimates from fixed effect maximum likelihood estimation (MLE) are reported with robust standard errors in parentheses. The sample size is 21,291 (3,043 counties for each year from 1933 to 1939 excluding 10 observations with no births in the year). Column (1) replicates our baseline estimates from Column (6) of Table 4. Column (2) includes other county controls interacted with year fixed effects rather than a linear trend. Column (3) extends the sample to 1941. Column (4) excludes those counties which border Canada or Mexico (n = 113). The resulting sample size is 20,500 (again, excluding 10 observations with no reported births in the year). Column (5) adds lead treatment variables for dryish and wet. Column (6) limits the sample to the period from 1933 to 1936. See Table 5 for the definition of each control variable. Significance levels: *** p<0.01, ** p<0.05, * p<0.1

Table 9: The Effect of Repeal on Infant Mortality Rates, Wet Counties versus Wet States

	(1)	(2)	(3)
Dryish in initial year	0.031***	0.031***	0.031***
	(0.012)	(0.012)	(0.012)
Dryish in subsequent years	0.027**	0.026**	0.027**
	(0.011)	(0.011)	(0.011)
Wet state in initial year	0.015	0.016	0.016
	(0.016)	(0.016)	(0.016)
Wet state in subsequent years	0.021*	0.024**	0.024**
	(0.012)	(0.012)	(0.012)
Wet county in initial year	0.003	0.003	0.003
	(0.008)	(0.008)	(0.008)
Wet county in subsequent years	0.017*	0.017*	0.018**
	(0.009)	(0.009)	(0.009)
Log likelihood	-62,589	-62,581	-62,580
N	21,291	21,291	21,291
N of county	3,043	3,043	3,043
County FEs	X	X	X
Year FEs	X	X	X
Retail sales	X	X	X
Other county controls	X	X	X
New Deal spending		X	X
Medical institutions			X
Hospital beds			X

Notes: Estimates from fixed effect maximum likelihood estimation (MLE) are reported with robust standard errors in parentheses. The sample size is 21,291 (3,043 counties for each year from 1933 to 1939 excluding 10 observations with no births in the year). Here, wet is divided into those counties which went wet through local option (*wet county*) and those which went wet through state legislation (*wet state*). The latter include: Arizona, California, Indiana, Nevada, and South Dakota in 1934; Delaware, Idaho, Iowa, Montana, South Carolina, Utah, and Wyoming in 1935; and North Dakota in 1937. Out of 2,180 ever-wet counties, 600 are wet-state, and 1580 are wet-county.See Table 5 for the definition of each control variable. Significance levels: *** $p<0.01$, ** $p<0.05$, * $p<0.1$

Appendix A: An Illustrative Model of Alcohol Consumption

Suppose that individuals have a willingness to pay for alcohol, v, drawn from a county-specific (time-invariant) distribution, $F_c(v)$. Suppose that the cost of buying alcohol under prohibition (that is, in a dry county) is d and the cost when buying without prohibition restrictions (that is, in a wet county) is w and that $d > w$. Suppose that each individual faces a binary choice of whether or not to buy alcohol based on whether their v exceeds the cost of purchase. Then, the purchase rate is $1 - F_c(d)$ in a dry county and $1 - F_c(w)$ in a wet county. The causal effect of ending prohibition on alcohol purchases is thus $1 - F_c(w) - (1 - F_c(d)) = F_c(d) - F_c(w)$.

Suppose that individuals prefer to change the law to end prohibition if for some cutoff $k \leq w$, their willingness to pay satisfies $v > k$. This cutoff k is assumed to be less than or equal to w because individuals who value alcohol at least enough to buy it when it is legal will be interested in ending prohibition. Additionally, some individuals who would not buy it might value freedom of choice as a generic good or might value the reduced criminal activity associated with ending prohibition. Then, under majority rule, prohibition will end in those counties where $1 - F_c(k) > 0.5$, or, equivalently, where $F_c(k) < 0.5$. Thus, the distribution F_c determines both whether or not a county will end prohibition *and* the causal effect on alcohol consumption. This implies that we have both heterogeneous treatment effects and correlated heterogeneity determining which counties are treated.

Since the only heterogeneity we allow for in this simple model is time-invariant, the inclusion of county fixed effects is enough to ensure that treatment is exogenous to the regressors. This implies that we can obtain an unbiased estimate of the average treatment effect on the treated (ATT) for wet counties. However, since the model tells us that treatment effects are heterogeneous, the ATT for wet counties does not in general equal the average treatment effect (ATE) for the population.

Individuals who have a willingness to pay greater than or equal to d are *drinkers*, because they will buy alcohol whether or not it is prohibited. Individuals who have willingness to pay less than or equal to zero are *abstainers*, because they will not buy alcohol at any price. The normal distribution implies that all counties have a nonzero fraction of abstainers who do not value alcohol consumption and a nonzero fraction of drinkers who highly value alcohol consumption. Neither abstainers nor drinkers change their binary alcohol purchase decision in response to changes in legal prohibition.

Suppose that F_c is normal with unit variance and heterogeneous medians, m_c, which vary across counties. Counties whose county-specific median, m_c, is low would have nearly all abstainers and counties whose median is high would have nearly all drinkers. A county with mostly drinkers would have a very high m_c so that both w and d would be very far below the median. Consequently, such a county would go wet, but the causal effect on alcohol purchase will be small. Such counties have high fractions of the population purchasing alcohol regardless of its prohibition status.

A county with mostly abstainers would have a very low median so that have both w and d would be very far above the median. Such a county would stay dry, but if it went wet, it would also have a very small causal effect on alcohol consumption.

Counties with m_c just below w might go either way on staying dry, but those that become wet would tend to have a large causal effect because the density of a normal is highest at the median. The observed ATT will, thus, depend on the distribution of m_c across counties. If the empirical distribution of m_c across counties is right-skewed, for example, the ATT could be quite small because the counties full of drinkers would dominate the counties that ended prohibition.

Now consider an intermediate treatment: having a wet neighbor but remaining dry (that is, dryish in the language of the main text). This reduces the cost of purchase from d to n, where $w < n < d$ gives the cost of buying alcohol from a neighboring county, which is less than the cost of buying it locally under prohibition, but more than the cost of buying it in a wet county. The source of this price friction is at least twofold: (1) travel distance between one's residence and the county line of the nearest wet neighbor and (2) potential local enforcement of prohibition restrictions in the home county.

The treatment effect for dryish is then $F_c(n)\text{-} F_c(w)$. Conditional on the distribution of willingness-to-pay in the country, this treatment effect is strictly smaller than that of transitioning to wet status, because we integrate a smaller range of the willingness-to-pay distribution. Further, only a dry county can be treated with this intermediate treatment. For these two reasons, we argue that the treatment effect of a dry county having a wet neighbor provides a *conservative* (lower-bound) estimate of the ATT for a dry county becoming wet. This is an interesting object for two reasons: (1) we can say something about the effect of treatment on agents that never choose treatment and (2) if we take our estimated ATT for wet as unbiased, then we can combine the estimates to get a lower-bound estimate of the ATE for the population.

The above arguments suggest that in this simple model, the treatment effect is non-monotonic in the county-specific preference parameters. If treatment is exogenous to the regressors (as in the fixed-effects estimate), then we can recover the ATT for wet counties. However, if treatment is endogenous, perhaps due to time-varying county-specific unobserved preference shocks, then one would typically appeal to instrumental variables to correct for that endogeneity. However, because the treatment effect is non-monotonic, standard instrumental variables approaches will not work.

We can also express this intuition more formally as a heterogenous treatment effects model. Let Y_{ct} be the alcohol purchase rate (binary for individuals, but a rate for a county-year). The above model implies that

$$Y_{ct} = T_{ct} * G_c + X_{ct} * B + \theta_c + \delta_t + e_{ct},$$

where T_{ct} is the treatment status (wet=1, dry=0) for county c in period t, X_{ct} are covariates, θ_c are county fixed effects, δ_t are year fixed effects, and e_{ct} are error terms. The important feature here is that the treatment effects G_c are heterogeneous across counties, so this a random coefficients model. Furthermore, there is an underlying parameter governing county-specific preferences for

57

consuming alcohol, m_c, that determines both T_{ct} and G_c. Under the model above, T_{ct} is monotonically increasing in m_c. Further, for m_c high enough, G is monotonically decreasing in m_c.

Given the restriction that variation in alcohol preferences (m_c) is time-invariant, then the only variable that treatment depends on is county. If we include county fixed effects as regressors, then we satisfy the selection-on-observables condition that guarantees that the estimated value of the coefficient on *wet* is an unbiased estimate of the average treatment effect on the treated (ATT) for counties that go wet.

Now assume that alcohol preferences vary over both time and county in ways uncorrelated with the observed covariates (so that we have m_{ct}). This implies

$$Y_{ct} = T_{ct} * G_{ct} + X_{ct} * B + \theta_c + \delta_t + e_{ct},$$

because m_{ct} causes variation in the treatment effect (G). Furthermore, assume that counties that go wet have m_{ct} high enough that the treatment effect (G) is monotonically decreasing in m_{ct}. In this case, including the county fixed effects does not deliver selection-on-observables, and the estimated coefficient on *wet* is a biased estimate of the ATT for counties that opt for wet status. However, because we know that T_{ct} increases in m_{ct} and G_{ct} decreases in m_{ct}, we can sign the bias: the estimated coefficient is a downward-biased estimate of the ATT for wet status.

Note that the channel for this form of endogeneity does not run through the additive error term (e_{ct}). Instead, the issue arises all through the fact that a single unobserved variable (alcohol preferences) drives both the response of alcohol consumption to a change in prohibition status and the probability of opting for a change in prohibition status.

Appendix B: A Comparison of Results under MLE and OLS

A very common specification in much of the literature is to regress the (logged) infant mortality rate on covariates of interest. Thus, we compare our MLE with unweighted and weighted OLS where the weight is the number of births in a county-year.

We begin with the observation that for small numbers of births B_{ct}, the binomial distribution of infants deaths D_{ct} is skewed and discrete, but as B_{ct} gets large, the distribution of infant deaths becomes more symmetric and smooth. As $B_{ct} \rightarrow \infty$, the distribution of D_{ct} is approximately normal by the de Moivre–Laplace theorem and is approximately distributed as $D_{ct} \sim N(B_{ct}\gamma_{ct}, B_{ct}\gamma_{ct}(1-\gamma_{ct}))$. Consequently, the infant mortality rate, $I_{ct} = D_{ct}/B_{ct}$, is approximately asymptotically distributed as $I_{ct} \sim N(\gamma_{ct}, \gamma_{ct}(1-\gamma_{ct})/B_{ct})$, leading to the widespread use of least squares estimates regressing the infant mortality rate on covariates. The limiting distribution of the binomial suggests that if the observed numbers of deaths are large, then least squares estimators are acceptable, but they should use weights B_{ct} to gain efficiency. That is, estimation should be by weighted least squares, given the dependence of the variance $\gamma_{ct}(1-\gamma_{ct})/B_{ct}$ on the observed numbers of births. Further, robust standard errors should be used to deal with the fact that the variance also depends on the unobserved observation-specific γ_{ct}.

Following the same notation from the fixed effects binomial model in the main text, we estimate the following OLS equation:

$$log(D_{ct} / B_{ct}) = \beta_0 + \beta_1 W_{ct} + \beta_2 N_{ct} + \beta_3 W_{ct*} + \beta_4 N_{ct*} + \gamma X'_{ct} + \theta_c + \delta_t + \varepsilon_{ct}$$

where D_{ct} and B_{ct} are the number of infant deaths and the number of births at county c in time t. W_{ct} and N_{ct} are equal to one in all the years following the change in prohibition status and equal to zero in the year of—and all years preceding—the change in prohibition status for wet and neighboring counties, respectively. Similarly, W_{ct*} and N_{ct*} are equal to one in the year of the status change and equal to 0 in all other years to allow for partial treatment effects in the year of change in prohibition status. The control group is then the set of bone dry counties, dry counties without any wet neighbors. X_{ct} is a set of the time-varying county characteristics which is the same as in our baseline estimate found in column (6) of Table 4. λ_c and θ_t are county and time fixed effects, respectively. Our coefficients of interests are β_2 and also potentially β_1.

This specification presents another issue which arises with OLS: observations with no infant deaths must be dropped or scaled before taking the log. Following the convention in the literature (e.g., Carpenter and Dobkin, 2009), we add 0.5 to the value of infant deaths in order to retain the 342 observations that have no reported infant deaths. Unweighted OLS estimates do not change if instead we add 0.1.

For ease of comparison, columns (1) through (3) in Table B1 replicates the MLE estimates from columns (4) through (6) in Table 4. Columns (4) through (6) present the results of an equivalent regression of the logged infant mortality rate on the same covariates estimated via unweighted OLS. Comparing our preferred results in column (3) to those in column (6), we see that unweighted OLS fails to detect any statistically significant effects associated with *dryish in initial year*. We also see that the estimated effect of *dryish in subsequent years* in column (6) is

nearly twice as large as that in column (3). Likewise, we see a 50% larger estimated effect associated with *wet in subsequent years* in column (6) as opposed to column (3). This suggests that unweighted OLS may be problematic in that it may be placing undue weight on unrepresentative observations, e.g., particularly small counties.

Table B1: MLE versus Unweighted OLS Estimates

	MLE			Unweighted OLS		
	(1)	(2)	(3)	(4)	(5)	(6)
Dryish in initial year	0.038***	0.034***	0.031***	0.024	0.025	0.024
	(0.012)	(0.012)	(0.012)	(0.019)	(0.019)	(0.019)
Dryish in subsequent years	0.044***	0.038***	0.027**	0.053***	0.054***	0.052***
	(0.011)	(0.011)	(0.011)	(0.018)	(0.018)	(0.018)
Wet in initial year	0.013	0.013	0.005	0.020	0.020	0.019
	(0.009)	(0.009)	(0.008)	(0.013)	(0.013)	(0.013)
Wet in subsequent years	0.010	0.015	0.018**	0.026*	0.024*	0.028*
	(0.010)	(0.010)	(0.009)	(0.014)	(0.015)	(0.015)
Log likelihood/R-squared	-62,749	-62,725	-62,589	0.51	0.51	0.51
N	21,291	21,291	21,291	21,291	21,291	21,291
N of county	3,043	3,043	3,043	3,043	3,043	3,043
County FEs	X	X	X	X	X	X
Year FEs	X	X	X	X	X	X
Retail sales		X	X		X	X
Other county controls			X			X

Notes: Columns (1)–(3) report estimates from fixed effect maximum likelihood estimation (MLE) with robust standard errors in parentheses. Columns (4)–(6) report the estimates from unweighted OLS with standard errors clustered at county level in parentheses. The outcome in these columns is the logged infant mortality rate where we add 0.5 deaths to the 342 observations with no infant deaths before taking the log. Changing this number to 0.1 has no effects on our estimates. The sample size is 21,291 (3,043 counties times 7 years from 1933–1939 excluding 10 observations with no births in the year). See Table 5 for the definition of each control variable. Significance levels: *** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table B2 again compares our preferred set of estimates from MLE to those of an equivalent regression of the logged infant mortality rate on the same covariates, estimated this time via OLS weighted by the number of births in a county. Apart from the non-result on *wet in subsequent years* in column (6), the results are quite similar across the various columns. What is more, the R-squared of the regressions in columns (4) through (6) increases from Table B1 to Table B2. This suggests that OLS may generate reasonable results, provided that an appropriate set of weights can be defined and the arbitrary treatment of zeroes—whether by dropping or scaling them—is not pervasive.

Table B2: MLE versus Weighted OLS Estimates

	MLE			Weighted OLS		
	(1)	(2)	(3)	(4)	(5)	(6)
Dryish in initial year	0.038***	0.034***	0.031***	0.034**	0.031**	0.028**
	(0.012)	(0.012)	(0.012)	(0.013)	(0.013)	(0.013)
Dryish in subsequent years	0.044***	0.038***	0.027**	0.047***	0.043***	0.030**
	(0.011)	(0.011)	(0.011)	(0.013)	(0.013)	(0.013)
Wet in initial year	0.013	0.013	0.005	0.007	0.007	-0.002
	(0.009)	(0.009)	(0.008)	(0.008)	(0.008)	(0.008)
Wet in subsequent years	0.010	0.015	0.018**	0.004	0.008	0.011
	(0.010)	(0.010)	(0.009)	(0.010)	(0.010)	(0.010)
Log likelihood/R-squared	-62,749	-62,725	-62,589	0.69	0.69	0.70
N	21,291	21,291	21,291	21,291	21,291	21,291
N of county	3,043	3,043	3,043	3,043	3,043	3,043
County FEs	X	X	X	X	X	X
Year FEs	X	X	X	X	X	X
Retail sales		X	X		X	X
Other county controls			X			X

Notes: Columns (1)–(3) report estimates from fixed effect maximum likelihood estimation (MLE) with robust standard errors in parentheses. Columns (4)–(6) report the estimates from OLS with standard errors clustered at county level in parentheses, weighted by the number of births in a county. The outcome in these columns is the logged infant mortality rate where we add 0.5 deaths to the 342 observations with no infant deaths before taking the log. Changing this number to 0.1 has no effects on our estimates. The sample size is 21,291 (3,043 counties times 7 years from 1933–1939 excluding 10 observations with no births in the year). See Table 5 for the definition of each control variable. Significance levels: *** p<0.01, ** p<0.05, * p<0.1

## Appendix C: Incorporating the Spatial Distribution of Counties into the Empirical Model

In the results presented above, our main explanatory variable of interest is dryish, an indicator variable for whether a dry county has at least one wet neighbor. One of the primary reasons for using this particular specification is to increase the statistical power of our estimation. However, dryish counties may be heterogeneous. In particular, the spatial distribution of counties might need to be taken into account. That is, what might matter is not only whether a neighboring county opts for wet status, but also how far away that county is. For instance, counties in the eastern and western halves of the United States have widely divergent sizes: New York County comprises 59 square kilometers while San Bernardino County comprises 52,072 square kilometers, implying that their respective neighbors could be very near or very far. Figure C1 below lays out the cumulative distribution of dryish county seats by distance to the border of the nearest wet county. The sample is comprised of the 700 counties which are dryish at some point in between 1934 and 1939. The horizontal line corresponds to the first quartile of 13.0 kilometers, suggesting dryish counties were, in the main, close to their wet neighbors (for presentation purposes, we limit the value of the *x*-axis to 80 kilometers while the maximum value observed in sample is 125.8 kilometers).

Figure C1: Cumulative Distribution of Dryish Counties by Distance to Nearest Wet County
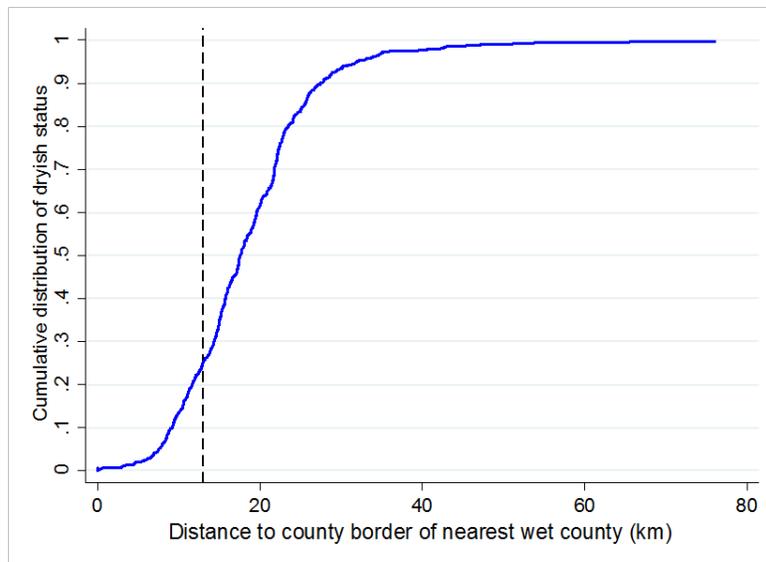
As a first step, we can consider the unconditional relationship between distance to the nearest wet county and infant mortality rates as in Figure C2. More specifically, the raw correlation between the distance to the county border of the nearest wet county and the logged infant mortality rate is plotted. The sample is comprised of the 700 counties which are dryish at some point in between 1934 and 1939 ($N = 2,546$). We use the command "binscatter" in Stata (Stepner, 2016) to generate an equal-sized binned scatterplot. The solid line is the line fitted by OLS when we assign the distance in the contemporaneous year for the initial year when the county becomes dryish and the distance from the previous year for the subsequent years, in order to be consistent with our main specification. Thus, there is some indication of a negative relationship in between the two variables.

Figure C2: Distance to Nearest Wet County and Infant Mortality Rate
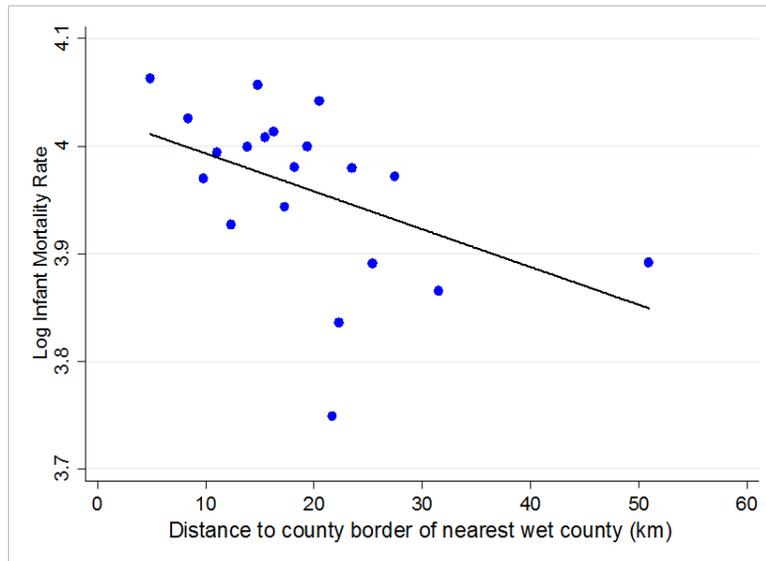
We now formalize this relationship by incorporating distance to the nearest wet county into our empirical model and replicating the specifications of Table 5. Here, we make a distinction between dryish counties which are near and dryish counties which are far from neighboring wet counties. Ideally, we would like to split this distance to the nearest wet county into multiple bins in order to estimate the differential effects by distance in a flexible way (e.g., Lovenheim and Slemrod, 2010). However, since the number of dryish counties is not very large, we will unfortunately lack power in that type of specification. We instead choose to impose a degree of parametrization to increase the statistical power. In this case, we would like our distance measure to satisfy the following two conditions: (1) its value should approach one as the distance to the nearest wet county approaches zero and (2) its value should approach zero as the distance to the nearest wet county increases.

One natural starting point might be the inverse of the distance to the nearest wet county. However, its value approaches infinity as the distance approaches zero. Slightly modifying this measure, we define an inverse-distance measure as follows:

$$\text{Inverse distance} = \begin{cases} 1 & \text{if } d \leq L \\ L/d & \text{if } d > L \end{cases}$$

where d is the distance to the nearest wet county and the L is a threshold value. In this way, our *inverse distance* measure is bounded between 0 and 1. To be consistent with our main specification, we assign the distance (d) in the contemporaneous year for the initial year and assign the distance from the previous year for the subsequent years. In terms of the distance measure (d), we consider both the distance from a county seat to the border of the nearest wet county (our preferred choice) and the distance from a county seat to the county seat of the nearest wet county. In terms of the threshold distance (L), we take the 25th percentile of distance to the nearest wet county among dryish county as our default. We also report the results using the 5th and 10th percentiles. Thus, our default L, the 25th percentile value for the distance to the

nearest wet county border, is 13.0 kilometers as shown in Figure C1 while that for the distance to the nearest wet county seat is 28.6 kilometers.

To follow our main specification, we also separately include wet in the initial year and wet in subsequent years. Thus, the control group is bone-dry. In this way, we exploit the heterogeneity among dryish counties where the variation comes from the distance to the nearest wet county. However, we choose to not take any of these specifications as our baseline because:
(1) there is a relative lack of power when we incorporate this heterogeneity in distance; (2) the inverse distance measure likely suffers from measurement error, primarily for the fact that we do not know the distribution of population within the counties and, thereby, cannot calculate population-weighted distances to the nearest wet county; and (3) it is unclear what functional form any inverse distance measure should take as the decay rate of a variable like dryish is unknown.

With these caveats in mind, Table C1 presents the results of our preferred specification incorporating distance to the border of the nearest wet county. Table C2 presents the results of a very similar specification which incorporate distance to the nearest wet county seat while Table C3 presents the results for both measures of distance, but with varying threshold values (*L*).

Consider column (1) of Table C1 as it is the direct analog of our baseline results in column (6) of Table 4. Again, we have statistically significant effects estimated for both dryish counties in initial and subsequent years along with *wet in subsequent years*. Naturally though, the interpretation of the coefficients is slightly different. Here, the value of 0.031 for *inverse distance in subsequent years* suggests that dryish counties which were 13 kilometers or less away from a wet border experienced a 3.1% rise in infant mortality while dryish counties which were more than 13 kilometers from a wet border experienced less than a 3.1% rise in infant mortality. For instance, moving from the 25th percentile of distance (13.0 kilometers) to the 75th percentile of distance (22.3 kilometers) entailed a decline in the estimated effect of being dryish from 3.1% to 1.8%.

The remaining columns of Table C1 suggest that materially the same results arise across the use of different controls. Table C2 speaks to a broad equivalence of results when using the distance separating dryish county seats from wet county seats as opposed to wet county borders, although the results for wet county borders generally have greater power. Finally, Table C3 suggests that our distance-related results are robust, regardless of the choice of the threshold value. Cumulatively, the tables speak to the fact that geography was an important factor in mediating our results.

Table C1: The Effect of Repeal on Infant Mortality Rates, Distance from Dryish County Seats to Nearest Wet County Borders

	(1)	(2)	(3)	(4)	(5)
Inverse distance in initial year	0.043***	0.042***	0.043***	0.041***	0.031**
	(0.014)	(0.014)	(0.014)	(0.014)	(0.015)
Inverse distance in sub. years	0.031**	0.030**	0.030**	0.028*	0.027
	(0.014)	(0.014)	(0.014)	(0.014)	(0.017)
Wet in initial year	0.005	0.005	0.005	0.003	0.012
	(0.008)	(0.008)	(0.008)	(0.008)	(0.015)
Wet in subsequent years	0.017**	0.018**	0.018**	0.012	-0.002
	(0.008)	(0.008)	(0.009)	(0.009)	(0.018)
Log likelihood	-62,588	-62,581	-62,580	-62,445	-61,860
N	21,291	21,291	21,291	21,291	21,291
N of county	3,043	3,043	3,043	3,043	3,043
County FEs	X	X	X	X	X
Year FEs	X	X	X	X	X
Retail sales	X	X	X	X	X
Other county controls	X	X	X	X	X
New Deal spending		X	X	X	X
Medical institutions			X	X	X
Hospital beds			X	X	X
State-trends				X	
State-year FEs					X

Notes: Estimates from fixed effect maximum likelihood estimation (MLE) are reported with robust standard errors in parentheses. The sample size is 21,291 (3,043 counties for each year from 1933 to 1939 excluding 10 observations with no births in the year). *Inverse distance* takes a value of one if the distance from a dryish county seat to the nearest wet county border (d) is less than or equal to the threshold (L) and takes the value of L/d if the distance is greater than the threshold (L). Therefore, *inverse distance* is bounded between 0 and 1. To be consistent with our main specification, we assign the distance (d) in the contemporaneous year for the initial year and assign the distance from the previous year for the subsequent years. Here, we take the 25th percentile of distance to the nearest wet county border which is 13.0 kilometers (as shown in Figure C1). Column (1) includes our baseline covariates. *Retail sales* are county-level retail sales per capita which is used as a proxy for income. Other county controls in column (2) are the variables reported in columns (1) and (3) of Table 3 interacted with a linear trend, except for *retail sales* which is time-varying. *New Deal spending* in column (2) is the cumulative amount of county-level New Deal spending per capita interacted with a linear trend. *Medical institutions* in column (3) is the number of medical institutions per 1000 people in a county while *hospital beds* is the number of hospital beds per 1000 women aged 15–44 in a county (both of which are time-varying). Significance levels: *** p<0.01, ** p<0.05, * p<0.1

Table C2: The Effect of Repeal on Infant Mortality Rates,
Distance from Dryish County Seats to Nearest Wet County Seats

	(1)	(2)	(3)	(4)	(5)
Inverse distance in initial year	0.044***	0.043***	0.043***	0.042***	0.030**
	(0.015)	(0.015)	(0.015)	(0.015)	(0.015)
Inverse distance in sub. years	0.028*	0.027*	0.027*	0.025*	0.022
	(0.014)	(0.014)	(0.014)	(0.014)	(0.017)
Wet in initial year	0.005	0.005	0.005	0.003	0.011
	(0.008)	(0.008)	(0.008)	(0.008)	(0.015)
Wet in subsequent years	0.016*	0.017**	0.017**	0.012	-0.005
	(0.009)	(0.009)	(0.009)	(0.009)	(0.018)
Log likelihood	-62,588	-62,581	-62,580	-62,444	-61,860
N	21,291	21,291	21,291	21,291	21,291
N of county	3,043	3,043	3,043	3,043	3,043
County FEs	X	X	X	X	X
Year FEs	X	X	X	X	X
Retail sales	X	X	X	X	X
Other county controls	X	X	X	X	X
New Deal spending		X	X	X	X
Medical institutions			X	X	X
Hospital beds			X	X	X
State-trends				X	
State-year FEs					X

Notes: Estimates from fixed effect maximum likelihood estimation (MLE) are reported with robust standard errors in parentheses. The sample size is 21,291 (3,043 counties for each year from 1933 to 1939 excluding 10 observations with no births in the year). *Inverse distance* takes a value of one if the distance from a dryish county seat to the nearest wet county border (d) is less than or equal to the threshold (L) and takes the value of L/d if the distance is greater than the threshold (L). Therefore, *inverse distance* is bounded between 0 and 1. To be consistent with our main specification, we assign the distance (d) in the contemporaneous year for the initial year and assign the distance from the previous year for the subsequent years. Here, we take the 25th percentile of distance to the nearest wet county seat which is 28.6 kilometers. Column (1) includes our baseline covariates. "Retail sales" are county-level retail sales per capita which is used as a proxy for income. Other county controls are the variables reported in columns (1) and (3) of Table 3 interacted with a linear trend, except for "Retail sales" which is time-varying. *New Deal spending* in column (2) is the cumulative amount of county-level New Deal spending per capita interacted with a linear trend. *Medical institutions* in column (3) is the number of medical institutions per 1000 people in a county while *hospital beds* is the number of hospital beds per 1000 women aged 15–44 in a county (both of which are time-varying). Significance levels: *** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table C3: The Effect of Repeal on Infant Mortality Rates,
Distance from Dryish County Seats to Nearest Wet Counties Using Different Thresholds

	Nearest wet county borders			Nearest wet county seats		
Percentile of threshold (L)	5th	10th	25th	5th	10th	25th
Threshold (L) in km	7.5	9.2	13.0	16.0	20.9	28.6
	(1)	(2)	(3)	(4)	(5)	(6)
Inverse distance in initial year	0.059***	0.052***	0.043***	0.065***	0.054***	0.044***
	(0.020)	(0.017)	(0.014)	(0.021)	(0.017)	(0.015)
Inverse distance in sub. years	0.042**	0.034**	0.031**	0.028	0.029*	0.028*
	(0.020)	(0.017)	(0.014)	(0.022)	(0.018)	(0.014)
Wet in initial year	0.005	0.005	0.005	0.004	0.005	0.005
	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)
Wet in subsequent years	0.016*	0.016*	0.017**	0.014*	0.015*	0.016*
	(0.008)	(0.008)	(0.008)	(0.008)	(0.009)	(0.009)
Log likelihood	-62,589	-62,589	-62,588	-62,588	-62,588	-62,588
N	21,291	21,291	21,291	21,291	21,291	21,291
N of county	3,043	3,043	3,043	3,043	3,043	3,043
County FEs	X	X	X	X	X	X
Year FEs	X	X	X	X	X	X
Retail sales	X	X	X	X	X	X
Other county controls	X	X	X	X	X	X

Notes: Estimates from fixed effect maximum likelihood estimation (MLE) are reported with robust standard errors in parentheses. The sample size is 21,291 (3,043 counties for each year from 1933 to 1939 excluding 10 observations with no births in the year). *Inverse distance* takes a value of one if the distance from a dryish county seat to the nearest wet county border (d) is less than or equal to the threshold (L) and takes the value of L/d if the distance is greater than the threshold (L). Therefore, *inverse distance* is bounded between 0 and 1. To be consistent with our main specification, we assign the distance (d) in the contemporaneous year for the initial year and assign the distance from the previous year for the subsequent years. For columns (1) through (3), we change the threshold (L) to the 5th, 10th, and 25th percentile of distance to nearest wet county border, respectively. Thus, column (3) replicates column (1) of Table C1. Columns (4)–(6) instead use the distance to nearest wet county seat with column (6) replicating column (1) in Table C2. *Retail sales* are county-level retail sales per capita which is used as a proxy for income. Other county controls are the variables reported in columns (1) and (3) of Table 3 interacted with a linear trend, except for *retail sales* which is time-varying. Significance levels: *** p<0.01, ** p<0.05, * p<0.1