J O I N T C E N T E R

AEI-BROOKINGS JOINT CENTER FOR REGULATORY STUDIES

Population and Regulation

Casey B. Mulligan and Andrei Shleifer

Related Publications 04-04
March 2004

We appreciate the comments of Lucian Bebchuk, Gary Becker, Dora Costa, Simeon Djankov, Ed Glaeser, Claudia Goldin, Louis Kaplow, Jim Hines, Larry Katz, Matt Kahn, Sam Peltzman, seminar participants at Clemson, Harvard, and the University of Chicago, data from Steve Tenn, and the research assistance of Herbert Lian, Pakshun Ng, Walter Theseira, and Vasu Vadlamudi. Mulligan gratefully acknowledges the financial support of the University of Chicago's Stigler Center for the study of the Economy and the State, and computing support from the NBER.

Executive Summary

We present a model of efficient regulation along the lines of Demsetz (1967). In this model, setting up and running regulatory institutions takes a fixed cost, and therefore jurisdictions with larger populations affected by a given regulation are more likely to have them. Consistent with the model, we find that higher population U.S. states have more pages of legislation and adopt particular laws earlier in their history. We also find that specific types of regulation, including the regulation of entry, the regulation of labor, and the military draft are more extensive in countries with larger populations. Overall, the data show that population is an empirically important determinant of regulation.

Population and Regulation

Casey B. Mulligan and Andrei Shleifer

I. Introduction

In a classic paper, Demsetz (1967) argues that the creation of institutions is shaped by demand. Introducing an institution only becomes efficient when the scale of an activity it supports becomes significant enough to cover the costs of creating and running it. Using the example of Indians in the Quebec region circa 1700, Demsetz maintains that the aggregate value of fur trading explains the emergence of enforced land ownership rights. In this paper, we show that Demsetz's logic is quite general theoretically but also valid empirically. We show that population of a community – our measure of the scale of various activities – is an important determinant of whether and how much these activities are regulated by that community.

The focus on population as a determinant of regulation is uncommon in theoretical and empirical work. Much of research on regulation has focused on contrasting public interest theories, in which regulation cures market failure (e.g., Pigou 1934), with public choice theories, in which regulation helps special interests to make money (e.g., Tullock 1967, Stigler 1971, Peltzman 1976, or McChesney 1987). This work generally focuses on particular industries to understand the structure of regulation. Our work is closer to the recent writings on regulation that see it as one strategy of enforcing good conduct, which needs to be compared to private orderings, dispute resolution in courts, and other enforcement strategies (e.g., Glaeser, Johnson and Shleifer 2001, Djankov et al. 2003). However, that work ignores scale as a potentially important factor shaping public involvement. Our purpose is to explore the implications of fixed organization, adoption, and enforcement costs of regulation on its use by focusing on total population of a community – a U.S. state or a country – as a determinant of its regulatory structure.

Section II presents a simple model of the creation of regulation of different types. The model establishes the relationship between population and regulation, but also permits a calculation of the population elasticity of regulation. Section III presents our cross-state data on the amount of law, showing how bigger states regulate more activities and tend to regulate each activity in more detail. Section IV revisits the previous literature on regulatory adoption, adding population to the list of explanatory variables, and shows how larger states tend to adopt regulations earlier. Section V looks at five types of regulation across countries, showing how high population countries tend to regulate a given activity more. Section VI extends the basic model to consider redistributive and not just efficient regulation, shows that many of the same insights apply, and presents some evidence on the empirical significance of that model. Section VII concludes.

II. Regulation Creation in a Simple Model with Fixed Costs and Zipf's Law

Our starting point is the efficiency approach to regulation which holds, in broadest terms, that a society chooses to regulate an activity when doing so is cheaper than alternative strategies of addressing social problems. Perhaps the simplest example of this approach, and one that we model, is resolution of disputes. Consider a jurisdiction, such as a U.S. state, where people interact with each other, and sometimes have a dispute. These disputes can be resolved in a number of ways. They can be resolved completely privately, perhaps because neighbors want to get along (Elickson 1991) or because participants in a given industry have established a private ordering to facilitate business in the long run (Greif 1993, Bernstein 1992). Alternatively, they can be resolved in court based on general principles of fairness, without any reference to legislation or regulation. Both of these strategies of enforcing good conduct have much to recommend them, but sometimes they fail. Private orderings are vulnerable to abuse of the weak by the strong; mafia enforcement is an example. Dispute resolution based on custom or common law is also expensive when application of general principles requires the assembly of a great deal of data and expertise. When private orderings and common law do not work well, mutually beneficial transactions might not take place.

As an efficient response to such problems, many societies introduce legislation and regulation, which delineate the rights and obligations of various parties in a community. Sometimes legislation just describes the rules of the game, and leaves the enforcement of these rules to private parties. Its main function is to reduce the cost of settling disputes in court. In other instances, enforcement is also taken over by the state, as in the case of regulatory agencies. At least in some instances, legislation and regulation reduce the transaction costs of enforcing good conduct relative to private orderings and common law. Glaeser, Johnson and Shleifer

(2001) argue, for example, that regulation reduces enforcement costs because regulators have stronger incentives to enforce rules than do the judges. Glaeser and Shleifer (2001) and LaPorta et al (2003) give examples of why "bright line rules" provided by regulation are easier to enforce – both publicly and privately – than torts. For our purposes, we do not assume that regulation is always more efficient, or identify the exact transaction costs – it is sufficient that, in some instances, regulation does reduce these costs.

As an example of this approach, consider the regulation of securities markets. In these markets, a crucial problem is the issuance of worthless securities to a poorly informed public by dishonest promoters and entrepreneurs. One can argue that underwriter and issuer reputation would eliminate this risk, but in many countries they do not; the temptation to take the money and run is just too strong. One can also argue that general principles of tort law, such as negligence, are sufficient to enforce good conduct and provide restitution to deceived investors. But in many cases, it is just too expensive to show that issuers and underwriters knew something material and negligently failed to disclose this information to investors. The questions about materiality of undisclosed information, of underwriter responsibility, and of issuer and underwriter intent are very expensive to answer. By delineating the exact duties of issuers and underwriters with respect to disclosure, and the exact burdens of proof on investors necessary to recover damages, a securities law can reduce the transaction costs of litigation, and thereby increase the level of activity in markets. The evidence provided by La Porta et al. (2003) is broadly supportive of this efficiency interpretation of securities laws and regulation.

Below we provide a formal model of regulation along the lines of this example, which helps us explore some of the implications of scale for the extent of regulation. But we stress that the focus on efficient dispute resolution, as in the example above, is only illustrative. First, regulation can in principle be a cost-reducing strategy for dealing with social problems in contexts other than just the resolution of disputes. As the size of a community rises, even if it remains very homogeneous in terms of income, education, or occupations, the variety of situations in which community members find themselves also rises. If a community wants to achieve some kind of uniformity of outcomes, it might resort to rules rather than markets or informal arrangements. Formulating and enforcing such rules takes a fixed cost, but they reduce the marginal cost of achieving social objectives.

Suppose, for example, that the issue is providing education for deaf children. In a small community, parents might get together and collectively decide how best to educate one or two deaf children. As the size of the community expands, so will the number of deaf children as well as the number of views their parents hold about the obligations of the community. The efficient way to resolve these disagreements in a larger community might be to establish formal rules for how deaf children are educated rather than to mediate community preferences in each case. As a second example, suppose that a country wants to raise an army. The market for volunteers is available regardless of country size, just as it is for the private sector. But the volunteer system may not fully achieve social objectives such as having an army representative of the nation's whole population, or minimizing the costs of tax collection. Once population becomes high enough, a country may find it worthwhile to pay the fixed costs of setting up and administering a draft system, so that it can turn away from the volunteer system for some of its military manpower needs.

Second, it is not important for our results whether regulations serve public rather than special interests, only that some component of the fixed organization, adoption, and enforcement costs are among the determinants of the extent of regulation. In our empirical work, we adopt this broader view of regulation, but in the model, we examine regulation as one a strategy for reducing the social cost of resolving disputes. Section VI begins to explore the question of whether the effects of population might be different if regulation serves special rather than public interests.

Let c denote the amount by which a dispute's resolution costs are reduced when there is legislation or regulation governing an activity. Interactions, and hence the nature of disputes, are heterogeneous. A day laborer's interaction, and potential disputes, with his employer are different from those between a salaried employee and his employer. We let $t \in [0,\infty)$ index the type of interaction that might occur in a community, or perhaps more literally the type of dispute that might occur. The index t is ordered so that the more frequent disputes have lower values of t. $f(\cdot)$ is a monotone decreasing density function, with $f(t)$ describing the likelihood that a randomly chosen dispute is of type *t*. When population is of size *N*, the total number of disputes of type *t* is $Nf(t)$.[1]

This specification assumes the same function *f* across all communities, and that the ordering of activities on the *t* axis is the same everywhere, which is clearly not true. Some communities specialize in particular economic activities or social groups. In this case, what determines the adoption of regulation is not the total population, but the number of people who would benefit (efficiently) from the activity being regulated. Moreover, communities might endogenously specialize in regulation, hoping to attract more of a given activity – as is the case with Delaware's specialization in corporate law. When we discuss empirical implications, we recognize that in some instances it is not the total population of the community, but rather the total affected population, that shapes the adoption of regulations.

Creating and enforcing regulations pertaining to a dispute of type *t* costs $r(t) = \rho + \beta Nf(t)$. This can be interpreted either as a social cost, or as a politically-weighted average cost. $r(t)$ has a fixed component ρ, and a variable component $\beta Nf(t)$, which is proportional to the total number of disputes of type *t*. For simplicity, we treat ρ, β, and *c* as constants even though in principle they can vary with GDP per capita or education.

Under the assumption of efficiency, legislation pertaining to disputes of type *t* is created *if and only if* the aggregate cost savings *c* exceed the aggregate regulation cost:

$$Nf(t)c \geq r(t) = \rho + \beta Nf(t) \qquad \textbf{(1)}$$

If $\beta \geq c$, there will be no regulation regardless of the dispute frequency or the population size. In particular, if $c < 0$, regulation is less efficient than private orderings or common law, and will not be adopted. Accordingly, we focus on the activities for which $\beta < c$.[2]

Since, by assumption, $f(\cdot)$ is monotone decreasing, equation (1) implies that there is a critical value *T* such that there is regulation pertaining to all disputes $t \leq T$, and no regulation pertaining to disputes $t > T$. Hence, *T* is the total range of regulation, and is determined by the formula:

---

[1]For simplicity, we have assumed that there is one dispute per capita. We could introduce a parameter indicating the number of disputes per capita, but it would just affect the measurement of the fixed regulation cost, rather then the regulation's population gradient.

[2]An interesting exercise in this case would be to think about forces that might lead to reductions in β. Becker and Mulligan (2003) and Peltzman (1989) conclude, for example, that regulation increases in response to its own enhanced efficiency.

$$T = f^{-1}\left(\frac{\rho}{(c - \beta)N}\right)$$

The fraction of disputes that are subject to regulation is simply *F*(*T*). Equation (2) says that regulation is increasing with population, and that the population elasticity is the same (in magnitude) as the fixed cost and variable cost elasticities.

The presence of fixed costs implies that jurisdictions that regulate an activity tend to have larger populations. In addition, if we fix a particular activity, such as the mining business, jurisdictions may vary according to exactly where mining is in their distribution *f* or what is the net variable cost (c-β) of not regulating. For example, mining may be a common activity in West Virginia, but uncommon in Georgia, so that if it were not for the fixed cost, a larger fraction *f* of West Virginia's population stands to benefit from mining regulation. In this example, *total affected population Nf*(*t*) determines regulatory adoption, so that some states like West Virginia may regulate mining even though their overall population *N* is small. More generally, some states with small overall populations will be among the states regulating a particular activity because of their large desired intensity (as measured by *f*, (c-β), or some combination).

Equation (2) says more in the interesting case where *f* is a particular member of the Pareto distribution family:[3]

$$f(t) = (t+1)^{-2} \quad \_ \quad T = \sqrt{\frac{(c-\beta)N}{\rho}} - 1$$

This Pareto distribution for disputes implies that the elasticity of the range of regulation *T* with respect to population is essentially constant and equal to 0.5. The fraction *F* of disputes that are subject to regulation has population elasticity $0.5F/(1-F)$, which can be greater or less than 0.5, depending on the (hard to measure) relative frequency of regulated and unregulated disputes.

One theoretical reason to expect the distribution of disputes to be Pareto with shape parameter near one derives from a simple "Gibrat's Law" model of the growth and decline of various forms of human interaction, by analogy with Gabaix' (1999) model of the growth of cities. Empirically, various forms of human interactions seem to occur at frequencies that are pretty nearly Pareto with shape parameter one. Examples include cities (Gabaix 1999), occupations,[4] homeowners insurance claims (Stuart 1983), and firm sizes (Axtel 2001).

III. <u>Regulation Across States Measured in KB's</u>

Population and the Amount of Law

One aggregate measure of regulation is the number of pages of law, made famous by Ronald Reagan in his first state of the union address when he recalled the reduction during his administration in the number of *Federal Register* pages. To compute this indicator of regulation for states, we measure the number of kilobytes (KBs) of unannotated *state* law in 37 states in 2001, 2002, or 2003. A kilobyte (KB) is 1024 bytes, and each byte represents a character. For example "Thou shalt not kill." is 20 bytes (including spaces and the period), or 0.0195 KB. We found that one page of law is roughly one kilobyte of law, and the typical state has tens of thousands of kilobytes.[5] Appendix I describes our algorithm for counting KBs of law for the states, and explains why 13 states were excluded from the counting.

Pages or KBs of law may measure the range of regulation T or, to the extent that the inframarginal disputes are regulated in more detail than the marginal ones, they may proxy for the fraction F of disputes that are subject to regulation. In either case, our model calibrated with Zipf's law says that the population elasticity should be about 0.5. Although not included in our model, there may be a "necessary" range of activities like murder, elections, or traffic, that are regulated in some detail regardless of population, in which case we might expect the population

[3]Equation (3) shows the regulation formula corresponding to a Pareto distribution f with shape parameter one. If the shape parameter were $\alpha\neq1$, then the exponent on $[(c-\beta)N/\rho]$ would be $1/(1+\alpha)$ rather than 0.5. Equation (3) applies only in the case where $\rho<(c-\beta)N$. Otherwise the fixed component of cost is too large and we simply have $T = 0$.

[4]We have calculated the 1990 Census PUMS distribution of workers by occupation, as 3-digit coded by the United States Census Bureau. There are 247823 employed men aged 25-54 with occupation coded. The most common code is 22 = " Managers and administrators, n.e.c.," for which there are 14874 observations (6.0% of the observations with occupation code). The second most common code is 804= "truck drivers," for which there are 11007 observations (4.4% of the sample). Overall, there are 191 occupation codes with a minimum of 249 observations, or 0.1% of the sample. In a graph (available upon request from the authors) of the occupation rank versus occupation frequency, using log scales, for the 191 occupations, we see nearly a straight line with slope -1.1, which means that the occupation types are distributed nearly Pareto with shape parameter near one.

[5]There are three reasons to measure KB rather than pages. First, not all states have their statutes published by the same publisher (or, within publisher, in the same format), so cross-state comparisons of pages would require adjustments for each publisher's font size, formatting style, etc. Second, some states have their total statutes

elasticity to be less than 0.5 and then rise with population to approach 0.5 as the regulation moves significantly beyond the necessary range. There is also a real possibility that a small state can adopt a regulation more cheaply by imitating the earlier-adopting large states. Imitation induces a positive correlation between population and the cost of adopting certain types of regulation. This positive correlation means that the cross-state population elasticity may be less than 0.5, especially among the smaller states whose statutes consist mainly of the "imitated" regulations.

The empirical relation between statute KB's and population is shown in Figure 1. The correlation (of the logs) is 0.88, and the overall regression elasticity is 0.31. The comparison of Delaware and Wyoming, and of Texas and New York, illustrates the basic fact. Delaware and Wyoming are similar in terms of total population, but different in terms of population density (Wyoming is the 2nd least densely populated state, while Delaware is 7th most). Given that both of these states have a similar number of statute KBs, and both fit near the regression line, it seems that population is much more important than population density in determining regulation. Texas and New York offer a similar comparison, since they have the same population, similar numbers of statute KBs, while New York is about 6 times more dense than Texas. Also notice that Delaware and New York state have similar population densities (almost 400 persons per square mile), but pages of statutes that differ by almost an order of magnitude, as predicted by our model.

Table 1 examines the validity and the robustness of the relationship between population and regulation illustrated in Figure 1. We include a southern state dummy, lawyers per capita, land area, income per capita, and several additional state characteristics as controls. The relationship between population and regulation remains strong, and the coefficient stays near .3. Southern states have fewer KB's of law. More law and more lawyers go together.[6] The regressions in Table 1 do not show any significant effect of income per capita on KB's of law. This is a surprising result, especially in light of the fact that Demsetz (1967) and North (1981) generally think of the aggregate level of economic activity as a determinant of adoption of

---

published irregularly, and publish only additions and retractions in the meantime. Third, bytes can be calculated (as described below) by computer and potentially be dis-aggregated by statute type.

[6]The number of lawyers is likely to be correlated with other determinants of number of laws. However, including this variable does not affect the estimated population elasticity because population and lawyers per capita happen to be uncorrelated. Appendix II explains how lawyers per capita are measured from the 1990 Census. Another

institutions. Indeed, our model's focus on population rather than aggregate activity distinguishes it from the earlier work. One reason that population rather than income may matter for regulation is that, as income rises, so do real wages and therefore the fixed labor cost of setting up and running a regulatory institution. If the fixed costs rise as fast as do the aggregate benefits per capita, population rather than income determines the adoption of a regulation.[7]

Alesina and Spolaore (2003) present a model of the determinants of country size, in which the benefits of spreading the fixed costs of a particular policy among more people are traded off against the inefficiency of implementing uniform policies in a heterogeneous population. In their view, population is positively correlated with heterogeneity such as ethnic diversity. In our model, population is a proxy for a different kind of heterogeneity – namely the likelihood of having at least some minimum number of people engaged in esoteric activities – which could be found in a large jurisdiction even if it were very homogeneous in terms of race, geography, education, or income. Empirically for U.S. states, does population proxy for heterogeneity and if so which kind? Specifications (1)-(3) suggest that population does not proxy for heterogeneity as measured by income inequality, occupational diversity, the importance of cities, or the prevalence of racial minorities, because these measures do not predict KB's or affect the estimated population elasticity.[8]

Figure 1 also does not readily tell us whether the amount of law in a state today depends on its current population or, since statutes accumulate over time, the population it had in the distant past and/or the number of years the state has existed (as a state). As shown in Table 1, these two variables (the former measured as 1920 population) have no power to predict statute KB's conditional on current population. We also include log of the number seats in each state's

measure of lawyers per capita for 2003 from the American Bar Association (2003) also does not help predict KB's of law, in part because it records (unlike the Census) NY and MA as extreme outliers.

[7]We are aware of only two studies that use absolute population as a determinant of regulation. Among the variables they use to predict adoption of "general incorporation code" by U.S. states, Shughart and Tollison (1985) include both aggregate population and aggregate manufacturing income (both in levels, with population measured in different years for different states). They find that populous states adopt later and states with more manufacturing income earlier, so it is hard to tell from their results whether populous states adopted earlier or later (we find the raw correlation between year of adoption and log 1910 population to be -0.55). McCormick and Tollison (1981) use absolute size of legislature to predict occupational licensing, and note that Stigler (1976) used absolute population to predict the size of legislature.

[8]See Appendix II for details on constructing the heterogeneity measures. We have also tried various measures of earnings inequality from the 1990 Census PUMS, and the fractions of income and employment in agriculture, with similar results.

House and Senate, in case the amount of law depends on the number of lawmakers which just happens to be correlated with population.[9] This control does not matter either.

Our model says that the population elasticity of regulation is 0.5, while the estimated elasticities shown in the first few columns of Table 1 are lower. We discussed above how in theory the elasticity may be smaller among the small states, and about 0.5 among the large states, because small states have regulations that are "necessary" or "imitable" from the big states. Figure 1 suggests, and specifications (4) and (5) confirm, that a larger elasticity in fact prevails among the larger states. Specification (4) is for the entire 37 state sample, and without the various controls other than Southern, and displays an elasticity of 0.33. Specification (5) throws out the 20 of the 37 states with below median population (i.e., states with a smaller 2000 population than Kentucky's 4,041,000), and displays an elasticity of 0.46. Although not shown in the Table, the estimated population elasticity is similar if we cut the sample at 3, 5, or 6 million rather then 4,041,000.

What Do KB's of Law Represent?

There are two reasons we believe that pages of statutes are correlated with the real amount of regulation. First, the aggregate time series of *Federal Register* pages (the *Federal Register* consists of new laws passed by Congress, executive orders, and federal government agency reports) deviates significantly from its trend during exactly those periods (since 1936) when it is commonly believed that federal regulation was accelerating the most – World War II and the 1970's. Second, as we show in Section IV, the population gradient seen in Figure 1 also appears in studies of the history of states' adoption of various laws, including occupational licensing, telegraph regulation, and worker's compensation. However, regulations like these diffuse pretty quickly from small to large states and, at least in the case of occupational licensing, the cross-state regulation-population gradient falls over time as regulation diffuses. Hence the pages we measure in 2003 may not have much to do with the adoption of regulations like occupational licensing and worker's compensation that began their diffusion decades ago, but rather with the adoption in more recent areas of regulation, with further elaboration of old regulations, or with the adoption of esoteric regulations by the big states that may never diffuse to the smaller ones.

---

[9]Log seats is more correlated with log 1920 population than with log 2000 population.

IL and IN are an interesting comparison, as the states are similar in many ways, except that IL has twice the population. Both states are near the regression line; IL has 40% more bytes of law than IN. Part of this difference is that IL has almost twice the bytes of criminal law and corrections.[10] Can these byte counts be attributed to a number of activities that are legal in IN and illegal in IL? Or do both prohibit the same set of activities and IL is just more detailed in its regulation of them? Several examples suggest that both differences are present.

Relative to IN, IL has many acts devoted to pretty minor issues (such as the "Coin Slug Act" and the "Peephole Installation Act").[11] Among the issues covered in the criminal law of both states, IL seems to regulate them in more detail. For example, IL has 359 KB devoted to drug offenses while IN has only 124 KB.[12] Included among IL's 359 KB is an entire "Drug Asset Forfeiture Procedure Act" (22 KB) devoted to the forfeiture of assets by persons involved in drug offenses, where IN has only a Civil Law chapter "Forfeiture of Property Used in Violation of Certain Criminal Statutes" (21 KB) on the seizing of assets of criminals, and this chapter applies to all kinds of criminals, including thieves, (media) pirates, smugglers, and terrorists.

Consider offenses related to animal fighting. IN has only a few sections (totaling 2 KB) in a chapter "Offenses Relating to Animals," while IL has a criminal section "Dog Fighting" plus two sections of the "Human Care for Animals Act" (with all three fighting sections totaling 11 KB). The IN statutes prohibit promotion, use of animals, or attendance with animal (or without) at animal fighting contests, and the possession of animal fighting paraphernalia. Conducting or attending a dog fight is also explicitly illegal in IL, but so is a whole range of other activities connected to dog fighting. Namely, IL explicitly prohibits a person to:

(a) "own, capture, breed, train, or lease a dog" for fighting,

(b) "promote, conduct, carry on, advertise, collect money for or in any other manner assist or aid in the presentation" of a dog fight,

---

[10]From the annotated statutes (not used for the KB counts cited in the text) we see that IL has many more annotations to its criminal law and corrections than does IN.

[11]IN covers coin slugs under a wider law regarding "Forgery, Fraud, and Other Deceptions" in reference to a slug that might be "deposited in a coin machine." IL's Coin Slug Act explicitly references "slug, washer, disc, token, string, cord or wire or by means of any false, counterfeited, mutilated, sweated or foreign coin, or by any means, method, trick, or device whatsoever not lawfully authorized by the owner of such coin box telephone, coin operated transit turnstile or transit fare box." In addition, IL has a "Telephone Coin Box Tampering Act."
To our knowledge, IN has no statute covering peephole installation (apartment units must be built with peepholes for the occupants to see out).

[12]Results are similar if we use statutes inclusive of annotations: IL has 426 drug pages while IN has 147.

(c) "sell or offer for sale, ship, transport, or otherwise move, or deliver or receive any dog which he or she knows has been captured, bred, or trained, or will be used, to fight another dog or human,"

(d) "manufacture for sale, shipment, transportation, or delivery any device or equipment which he or she knows or should know is intended for use in any [dog fight],"

(e) "possess, sell or offer for sale, ship, transport, or otherwise move any [dog fighting equipment],"

(f) "make available any site, structure, or facility, whether enclosed or not, that he or she knows is intended to be used for the purpose of conducting [a dog fight]".

IL law also details the procedures for sheltering animals found in connection with the enforcement of the animal fighting statutes.

The IN-IL comparison is likely to be representative of the population-animal regulation gradient for all 50 states. We have counted words of statute devoted to animal fighting for 37 states. Regressing log animal fighting words on log 2000 population and a dummy for south yields coefficients of 0.30 (s.e.=0.14) and 0.02 (s.e.=0.30), respectively.[13] The population elasticity for animal laws is the same as that for all laws combined.

IV. Evidence on Population and the Diffusion of Regulation Across States

Our model predicts that the regulatory difference between the more and the less populated states would be in the regulation of the interactions that are relatively infrequent, or that have relatively small net benefits. In the previous section, we have tested and confirmed this proposition using data on modern state statutes. Another approach is to look at the introduction of regulations, and to ask whether the more populous states introduce a given regulation earlier.

[13]Coefficient estimates are similar if we include proxies for the importance of agriculture.

Patterns of Adoption: Occupation and Industry Regulation

Stigler (1971) looks at the licensing of 37 occupations in the 48 mainland U.S. states. He predicts the year a state licenses an occupation using the prevalence and urbanization of that occupation in the state, and occupation fixed effects. Our model suggests that total population, or the *absolute* size of the occupation, should be added to the licensing year regressions. Roughly speaking, the difference between Stigler's specification and that suggested by our model is the inclusion of log total population as regressor in addition to, or instead of, occupational prevalence.[14] When we regress year of licensing on occupation dummies, the fraction of the population living in cities, and the log of 1910 population, the estimated population elasticity is -2.13 (standard error clustered by state = 0.55). Larger states tended to license occupations earlier.

To obtain an indicator of how pages of occupational regulation might have varied with population at various points in history, we transform Stigler's data set into repeated cross-sections. The year *y* cross-section lists states and the fraction of the 37 occupations it has licensed as of year *y*. The population elasticity is positive in the various cross-sections and, not surprisingly, is higher during the peak licensing years 1900-20. For example, the cross-state regression of fraction occupations licensed by 1850 on log 1950 population yields a coefficient of .005 (OLS standard error = .002). The regression coefficients are .031 (s.e.=.010) and .018 (s.e.=.015) if we measure occupations licensed as of 1910 and 1950, respectively.[15] This pattern suggests that the population-pages gradient we observe in 2002 derives more from regulations that are at the peak of their diffusion (or esoteric regulations adopted by big states that will never diffuse), than from regulations that mainly diffused years ago – like the licensing of Stigler's 37 occupations.

Other evidence is broadly consistent with our findings on Stigler's data. Nonnenmacher (2001) looks at the adoption of telegraph regulation circa 1850. Figure 2 graphs total state population, measured in 1850 and on a log scale, versus the year of first telegraph regulation for

[14]Stigler did not enter occupational prevalence (total number of persons practicing the occupation divided by total labor force) in log form, but if he had, and occupation size were normalized by population rather than labor force, then his specification and ours would differ only by a log population term. Stigler's estimated occupational prevalence coefficient was statistically insignificant. Another reason we emphasize log population, rather than log occupation size, as a regressor is that licensing may affect occupation size more readily than it affects population.

[15]Results are similar if we measure population in 1910 instead of 1950: the log population coefficients are .005, .031, and .011 for occupations licensed as of 1850, 1910, and 1950, respectively. Note that less than 1% of state-

each of the 32 U.S. states at the time. As expected, the populous states like NY, PA, MA, and VA were early adopters, and the last adopters (TX, FL, MN, IA, AR) were relatively unpopulated. The correlation between year of first law and log population is -0.56 (t-stat=3.73).

More populous states were also quicker to regulate working hours of women. Figure 3 graphs Landes' (1980, Table 1) report of the year of first maximum female working hours legislation against (log scaled) 1890 population. The correlation is -0.34 (t-stat=2.3). Among the 23 states for which Landes reports there being a minimum wage law for women, the correlation between year of first minimum wage law and log population is -0.29 (t-stat=1.0). TX, ND, DC, AR, and KS had an average 1890 population of 1.0 million and were the last of the 13 to legislate a minimum wage, while MA passed the first law and had an 1890 population of 2.2 million.

Population or Affected Population?

In Section II, we noted that the exact specification of the model might be too narrow, and that what matters for efficiency is the size of the population (and the number of interactions and conflicts) affected by a given regulation. Some evidence indeed supports this view.

Worker's compensation provides an interesting application of our analysis, because the population relevant for determining whether there will be worker's compensation regulation – namely, the individuals likely to be hurt in workplace accidents – can be quite different from the total state population, especially in the early 20th century when states were first taking up these laws. For example, the 1910 Census shows Wyoming's ranking 47th out of 48 in terms of total population, but 29th out of 48 in terms of total number of persons working as miners (a group likely to experience serious work injury). Figure 4 graphs *total* state miners found in the 1910 Census PUMS, on a log scale, against the year of first workers' compensation law (from Fishback and Kantor 2000, Table 4.3) for each of the 48 U.S. states at the time. The correlation is -0.27 (t-stat = 1.9). If we regress year of first law on log miners and miners per capita, log miners is the more important variable: the t-statistics on the regression coefficients are -1.5 and -0.5, respectively.

---

occupations were licensed as of 1850. The highest licensing rates were 1900-20, when more than one percent of the 1776 state-occupation cells became licensed each year.

The size of the affected population also predicts which states were early to pass legislation "prohibiting discrimination in employment on the grounds of race, creed, color, or national origin" (Landes 1968, p. 507). When Landes wrote, 18 northern states had laws *and* agencies to enforce them. As Table 2's "enforceable" column shows, the average 1950 black population in these states was 171,615, and three quarters of these states had at least 100,011 blacks. The four northern states with discrimination laws but no commission to enforce them (ID, ME, MT, VT) were much smaller – each had roughly 1,000 blacks. North Dakota and South Dakota were the only Northern states without laws; their black populations were 257 and 727, respectively. Table 2 is consistent with our hypothesis that small states are late to adopt regulation. However, in this case total population is correlated with population frequency; the last row of the Table shows that the regulating states were not only populous, but also had relatively large black population percentages.[16]

V. <u>Population and Regulation Across Countries</u>

A few studies have tried to measure regulations for a broad cross-section of countries. Here we briefly analyze the employment laws index of Botero et al (2003), the business entry regulation index of Djankov et al (2002), the death penalty measures of Mulligan et al (2004), and measures of military conscription. Table 3 reports one cross-country regression in each column. The columns differ according to their regulation measure (one of the four referenced above), and to whether a broader set of political variables (namely "left power") are included. The broadest sample is the 127 country "MGX" sample of Mulligan et al (2004), but includes only measures of death penalty, population, British legal origin, GDP per capita, democracy, whether a country belongs to Kornai's (1992) list of communist states, and some information about military activities. A narrower 71 country "LaPorta" sample of Botero et al. (2003) and Djankov et al. (2002) includes the regulation measures from those studies and a political measure of "left power".[17]

---

[16]Interestingly, among the 28 states with enforceable legislation, the correlation of year of first law with log total black population (black population share) is -0.26 (-0.11), respectively.
[17]Botero et al (2003) and Djankov et al (2002) have 85 countries. We exclude the former Soviet republics, Vietnam, and Lebanon, due to insufficient data on GDP for the years 1960-90.

Specifications (1)-(4) suggest that, holding constant legal origin, GDP per capita, democracy, and communism, more populous countries have more business entry procedures, employment regulations, and are more likely to have the death penalty. The results on business entry procedures and labor regulations are particularly supportive of our model since, generally speaking, incremental regulations in these areas cover more "issues" that might arise in the course of economic activity. Consistent with our model, a small jurisdiction might not think it worthwhile to deal with these issues through regulation, but a larger jurisdiction faces enough demand to cover the fixed costs.

Military conscription, as an addition to the volunteer army,[18] is a common form of reliance on regulation rather than contract to meet social objectives. Furthermore it is relatively easy to measure for a large panel of countries – on both the intensive and extensive margins. We obtain data on the existence of the draft, and on the number of draftees for 138 countries for the years 1985, 1990, and 1995 from *The Military Balance* published annually by the International Institute for Strategic Studies. Following Ross (1994), we use measures of the size of the Armed Forces, democracy, and economic development as predictors of a country's reliance on conscription.[19] Specifications (5) and (6) are like the previous four specifications, except that some military activity variables are added to the list of independent variables, and that the dependent variable is the fraction of years 1985, 90, and 95 enlisting conscripts in the armed forces. The common law effect, which is large and significant in all specifications, is broadly consistent with the general finding that, in labor as well as other markets, common law countries are more likely to rely on contract and less on regulation to achieve social goals (Djankov et al 2002, Botero et al. 2003).

The population effect is large, significant, and consistent with our theory in the sample of 127 countries. However, once we restrict attention to the subsample of 71 countries for which we have measures of left power, the size of the effect falls and its significance disappears. This is a consequence of using the subsample of richer and more democratic countries: the correlation between left power and log population is -.02 in the smaller sample, so omitting the left power

[18] To our knowledge, all militaries have some volunteers. For example, among the 68 countries in our sample having conscription (and reporting data on the number of draftees), the typical military force is about half volunteer and half drafted. The lowest volunteer intensities were in Senegal, Switzerland, Turkey, and Cyprus, where 5, 9, 15, and 15 percent, respectively, of their armed forces were volunteers. Hence, the question is not a volunteer versus a draft system, but whether to have a draft system supplement the volunteer system.

variable has essentially no effect on the left power coefficient. We also find that, when a country uses the draft, it uses it intensively, especially when the country is small (results not reported in the Table). Among the countries with less than median population and having a draft, 75% have more conscripts than volunteers. This observation is consistent with fixed costs of having a draft system, and inconsistent with the hypothesis that small countries just happen to intend to use the draft lightly.

VI. An Extension to the Politics of Regulation

We can extend the model to consider regulatory incidence. We suppose that each type of interaction involves two types of agents, whom we (arbitrarily) call "producers" and "consumers." For this purpose, we simplify our analysis by assuming that the fixed cost is entirely a cost of mobilizing an interest group to pass legislation, and that consumers and producers do not mobilize cooperatively. Each group faces a fixed cost of organizing, which we denote ρ and ρ' for producers and consumers respectively, and normalize group identities so that $\rho < \rho'$.[20] The social variable costs of regulating or not regulating are, as above, β and c, respectively. If only one of the groups organizes, regulation is passed and the organized group receives fraction $\theta > 0.5$ of the variable social surplus $(\beta\text{-}c)Nf(t)$. If both groups organize, they divide the variable surplus evenly.

The producers have the greater incentive to organize to push for regulation of an unregulated activity. Their criterion for organizing is an equation exactly like (1), except that the variable cost parameters β and c are multiplied by the sharing parameter θ. The total quantity of regulation T is then determined by an equation exactly like (2), except that the variable surplus term $(\beta\text{-}c)$ is multiplied by the sharing parameter θ, and we can derive all of the results cited above. Note that, when $\theta > 1$, the marginal regulations add to net social cost. For this reason, we refer to the regulation created by one party alone as "redistributive" and that created by both parties as "efficient," even though for some parameters even the "redistributive" regulation reduces social costs.

---

[19]We have also used government spending/GDP and the share of the population over the age of 65, and obtained similar results.

Because they have the larger organization cost, the consumers' organization decision involves not the tradeoff between regulation or not, but how the regulation surplus should be shared. When consumers are near the margin of being organized, the producers are already organized, so consumer organization serves the purpose of raising consumers' variable surplus from regulation by the amount $(\theta\text{-}0.5)(\beta\text{-}c)Nf(t)$, and lowering the variable surplus of the producers by the same amount. Consumers trade this benefit off against their organization cost ρ'. The scope of "efficient" regulation T_e is calculated in much the same way as we calculate T:

$$T_e = f^{-1}\left(\frac{\rho'}{(\theta - 0.5)\ (c - \beta)\ N} \right)$$

The quantity of redistributive regulation is $T\text{-}T_e$. $F(T)$ is, as above, the total fraction of interactions regulated, $F(T_e)$ is the fraction of interactions regulated in an efficient way, and the difference is the fraction of activities regulated in a redistributive way.

Both T and T_e increase with population. The ratio T_e/T can either rise or fall with population, depending on the distribution of interactions. With the shape-one Pareto distribution, T_e is essentially proportional to T, with the proportion determined by the relative fixed costs and the sharing parameter θ:

$$\frac{T_e + 1}{T + 1} = \sqrt{\frac{\rho}{\rho'}\ \frac{\theta - 0.5}{\theta}}$$

Figure 5 graphs formula (2) (amended to multiply $(c\text{-}\beta)$ by the sharing parameter θ) and the formula (4), showing how population, activity type, and the model parameters determine whether and how an activity is regulated. Activities are unregulated above the solid line and regulated below the solid line. The dashed line partitions the regulated activities into those regulated in an efficient vs. a redistributive way. Both schedules slope up, which implies three things. First, in a given jurisdiction, the infrequent activities are less likely to be regulated.

[20]The costs of political organization are not well understood, but one possibility is that producers organize more readily than consumers because the former is a more concentrated group and thereby suffer less from free-riding (Olson 1971).

Second, fixing the type of interaction (moving horizontally in Figure 5), smaller jurisdictions are less likely to have any regulation. Third, conditional on having a regulation of a given type (moving horizontally in Figure 5, in the region to the right of the solid schedule), the smaller jurisdictions are *more* likely to have redistributive regulations. The last result is intimately related to Becker's (1983) model of government policy, which holds that, when more interest groups with countervailing interests form, government policy becomes more efficient as a broader range of interests is articulated.

Do Small States have More Redistributive Regulation? The Case of Attorney Licenses

Our model distinguishes the margin of adopting a regulation (equation 2) from the margin determining whether a regulation is efficient or redistributive (equation 4). Small states are less likely to regulate a given activity but, conditional on regulating, *more* likely to regulate in a redistributive way. For some evidence on this point, we look at the regulation of lawyers. Because every state has licensing requirements for its lawyers, the operative margin is the degree to which those requirements redistribute income. Tenn (2002) has created, for 48 states and the three Census years 1970-90, an index of the degree to which licensing laws might raise the incomes of lawyers. He considers whether obtaining a license requires an examination, the difficulty of the examination, whether a lawyer must have a degree from an American Bar Association approved school, the residency requirements, and the amount of the bar exam fee, to form his index of licensing strictness, with stricter licensing delivering a higher value for the index. We calculate each state's time-averaged index, and graph it in Figure 6 versus log 1970 population. It appears that larger states have less strict regulations; the correlation is -0.27 (t-stat=-2.0).

VII. <u>Summary and Conclusions</u>

In this paper, we have presented a model of efficient regulation along the lines of Demsetz (1967). In this model, setting up and running regulatory institutions takes a fixed cost, and therefore jurisdictions with larger populations affected by a given regulation are more likely to have them. We then tested the model using data from both U.S. states and countries around the world. We found that higher population U.S. states have more pages of legislation and adopt particular laws earlier in their history. We also found that specific types of regulation, including the regulation of entry, the regulation of labor, and the military draft are more extensive in countries with larger populations. Overall, the data show that population is an empirically important determinant of regulation, consistent with the theory.

The results in this paper naturally raise the question of optimal regulatory jurisdictions. In both the efficiency and redistribution versions of our model, having a larger jurisdiction is better. In the efficiency model, regulation reduces the marginal cost of addressing social problems and so, with larger jurisdictions, a greater range of such problems is addressed efficiently. In the redistributive model of regulation, more interests are articulated when the jurisdiction is larger. However, in more general models, it no longer needs to be true that larger jurisdictions are always optimal. Competition among jurisdictions as in Tiebout (1956) or diversity of preferences among people (Alesina and Spolaore 2003) are arguments favoring smaller jurisdictions. Even with these considerations, however, our evidence suggests that the choice of an optimal jurisdiction must in part be determined by population.

Many of these results are most naturally interpreted from the efficiency perspective, but some evidence is also consistent with the public choice theories of regulation, in which the relevant fixed cost is that of organization rather than enforcement. The results also suggest that, because of increasing returns in regulation, we would expect to observe regulatory specialization, particularly in activities that can cheaply travel across jurisdictions. Delaware's specialization in corporate law is broadly consistent with the perspective of this model. Regardless of the exact model of fixed costs, the evidence is broadly supportive of the view that overcoming such costs is an important determinant of regulatory choice.

References

Alesina, Alberto and Enrico Spolaore. *The Size of Nations*. Cambridge, MA: M.I.T. Press, 2003.

American Bar Association. *National Lawyer Population by State*. Chicago: ABA Market Research Department, 2003.

Axtel, Robert L. "Zipf Distribution of U.S. Firm Sizes." *Science* 293, September 7, 2001: 1818-20.

Becker, Gary S. "A Theory of Competition among Pressure Groups for Political Influence." *Quarterly Journal of Economics*. 98(3), August 1983: 371-400.

Becker, Gary S. and Casey B. Mulligan. "Deadweight Costs and the Size of Government." forthcoming, *Journal of Law and Economics*. 2003.

Bernstein, Lisa. "Opting out of the Legal System: Extralegal Contractual Relations in the Diamond
Industry." *Journal of Legal Studies* 21, Month 1992, 115-157.

Botero, Juan, Djankov, Simeon, Rafael La Porta, Florencio Lopez de Silanes, and Andrei Shleifer. "The Regulation of Labor." NBER working paper no. 9756, June 2003.

Council of State Governments. *The Book of the States*. Chicago: Council of State Governments.

Demsetz, Harold. "Toward a Theory of Property Rights." *American Economic Review* 57(2), May 1967: 347-59.

Djankov, Simeon, Rafael La Porta, Florencio Lopez-de-Silanes, and Andrei Shleifer. "The Regulation of Entry." *Quarterly Journal of Economics* 117(1), February 2002: 1-37.

Djankov, Simeon, Edward Glaeser, Rafael La Porta, Florencio Lopez de Silanes, and Andrei Shleifer. "The New Comparative Economics." *Journal of Comparative Economics* 31, December 2003: 595-619.

Fishback, Price V., and Shawn Everett Kantor. *A Prelude to the Welfare State*. Chicago: University of Chicago Press for the NBER, 2000.

Gabaix, Xavier. "Zipf's Law for Cities: An Explanation." *Quarterly Journal of Economics* 114(3), August 1999: 739-67.

Glaeser, Edward L., Simon Johnson, and Andrei Shleifer. "Coase versus the Coasians." *Quarterly Journal of Economics* 116 (3), August 2001: 853-899.

Glaeser, Edward L. and Andrei Shleifer. "A Reason for Quantity Regulation." *American Economic Review* 91(2), May 2001: 431-35.

Goldin, Claudia. "Maximum Hours Legislation and Female Employment: A Reassessment." *Journal of Political Economy* 96(1), February 1988: 189-205.

Greif, Avner. "Contract Enforceability and Economic Institutions in Early Trade: the Maghribi Traders' Coalition." *American Economic Review* 83(3), June 1993: 525-48.

International Institute for Strategic Studies. *The Military Balance*. London, various issues.

International Monetary Fund. *Government Finance Statistics*. various issues.

Kornai, János. *The Socialist System: The Political Economy of Communism*. Princeton, NJ: Princeton University Press, 1992.

Landes, Elisabeth M. "The Effect of State Maximum-Hours Laws on the Employment of Women in 1920." *Journal of Political Economy* 88(3), June 1980: 476-94.

Landes, William M. "The Economics of Fair Employment Laws." *Journal of Political Economy* 76(4), July 1968: 507-52.

La Porta, Rafael, Florencio Lopez de Silanes, and Andrei Shleifer. "What Works in Securities Laws?" NBER Working paper #9882, August 2003.

Lee, Dwight R., and Richard B. McKenzie. "Reexamination of the Relative Efficiency of the Draft and the All-Volunteer Army." *Southern Economic Journal* 58(3), January 1992: 644-54.

McChesney, Fred S. "Rent Extraction and Rent Creation in the Economic-Theory of Regulation." *Journal of Legal Studies* 16(1), January 1987: 101-18.

McCormick, R.E. and R.D. Tollison. *Politicians, Legislation, and the Economy*. Boston: Martinus Nijhoff, 1981.

Mulligan, Casey B., Ricard Gil, and Xavier Sala-i-Martin. "Do Democracies Have Different Public Policies Than Nondemocracies?" forthcoming, *Journal of Economic Perspectives*, Winter 2004.

Nonnenmacher, Tomas. "State Promotion and Regulation of the Telegraph Industry, 1845-60." *Journal of Economic History* 61(1), March 2001: 19-36.

North, Douglass C., 1981, *Structure and Change in Economic History*, New York: Norton.

Olson, Mancur. *The Logic of Collective Action*. Cambridge, MA: Harvard University Press, 1971.

Peltzman, Sam. "Toward a More General Theory of Regulation." *Journal of Law and Economics* 19(2), 1976: 211-40.

Peltzman, Sam. "The Economic Theory of Regulation After a Decade of Deregulation." *Brookings Papers on Economic Activity*, Special Issue, 1989: 1-41.

Ross, Thomas W. "Raising an Army: A Positive Theory of Military Recruitment." *Journal of Law and Economics* 37(1), April 1994: 109-31.

Shughart, William F. and Robert D. Tollison. "Corporate Chartering: An Exploration of the Economics of Legal Change." *Economic Inquiry*. 23(4), October 1985: 585-99.

Stigler, George J. "The Sizes of Legislatures." *Journal of Legal Studies* 5, January 1976: 17-34.

Stigler, George J. "The Theory of Economic Regulation." *Bell Journal of Economics and Management Science* 2(1), 1971: 3-21.

Stuart, Charles. "Pareto-Optimal Deductibles in Property-Liability Insurance: the Case of Homeowner Insurance in Sweden." *Scandinavian Actuarial Journal* 9(4), 1983: 227-38.

Tenn, Steven. "Occupational Licensing in the Legal Profession." manuscript, November 2002.

Texas A&M University, Real Estate Center. "State Population by Decade." Online (http://recenter.tamu.edu/data/popsd/): Texas A&M University, 2002.

Tiebout, Charles M. "A Pure Theory of Local Expenditures." *Journal of Political Economy*. 64(5), October 1956: 416-424.

Tullock, Gordon. "The Welfare Costs of Tariffs, Monopolies, and Theft." *Western Economic Journal*. 5(3), 1967: 224-32.

United States Census Bureau. *Statistical Abstract of the United States*. Washington, DC: Census Bureau, 2001.

University of Virginia Geospatial and Statistical Data Center. *United States Historical Census Data Browser*. Online: University of Virginia, 1998 Available: http://fisher.lib.virginia.edu/census/.

Warner, John T., and Beth J. Asch. "The Economic Theory of a Military Draft Reconsidered." *Defence and Peace Economics* 7(4), 1996: 297-312.

Appendix I: Algorithm for Counting KB's of Law

Statute Types

There are two main formats for publishing state laws – "annotated" and "unannotated." The former are most commonly found in libraries – presumably because they are more useful to lawyers – and contain the text of each statute in effect at a point in time, *plus* some of the precedents that have affected interpretation of the statute and perhaps information about previous versions of the statute. The unannotated statutes contain only the text of the statute. Since we are interested in the relation between regulation and population, the distinction between annotated and unannotated statutes is important. A populous state is more likely to have had a court case that tested a particular statute, so we expect the quantity of annotations to increase with population.

Computer Algorithm for Counting KB's of Law

Every state has unannotated statutes available for browsing on the internet. The browsing is either in html, java, or pdf format, or in multiple formats. Our computer programs can only browse the internet in html format, so we were unable to make counts for 9 states which had laws on-line only in java format.[21] The byte counts of pdf files exceed the number of characters in the file (due to formatting), so we exclude the two states (Kentucky and North Dakota) for which on-line statutes are only in pdf format. The final two states, Georgia and New Hampshire, were excluded from the sample even though they had html statutes available, because they were not in a format accessible by our programs. As we show below, the 13 states excluded from our data set are very similar to the 37 included in terms of population and many other characteristics.

Statutes on-line are usually presented in a tree format, where users first browse a list of titles, for each title a list of chapters, etc., with the final nodes in the tree being the actual texts of laws. The tree format can be used to categorized formats by their legal classification, for example, tax, criminal, schools, occupations, or estates. States differ in terms of the number of levels in the tree, the number of final nodes used to present a given set of statutes, and hence the number of statutes and KB of statutes per final node. For example, New York has less than

[21]Those 9 states are Colorado, Kansas, Louisiana, Maryland, New Mexico, Ohio, Oklahoma, Virginia, Wisconsin. A number of states have both java and html statutes on line, including AR, CA, SD, TN.

4,000 final nodes, and South Dakota more than 40,000, even while the former has a lot more statutes than the latter. Since each html page usually includes headers and footers, this implies that the number of KB of html required to present a *given* set of statutes expands significantly with the number of html pages on which those statutes are presented. For example, SD has more bytes devoted to headers and footers than actual statutes, whereas NY has more than 80 bytes of actual statutes per byte of header or footer.[22] We therefore count statutes KB in four steps:

(1) A computer program automatically browses the entire html tree presenting a state's statutes, and downloads each www page from the tree, stripped of html tags. The statutes browsed were those in effect in 2001, 2002, or 2003.[23]

(2) A sample of downloads are visually inspected for a number of bytes of headers and footers on a typical html page.

(3) The number of html pages is multiplied by the result from step (2) to give total KB of headers and footers, and then subtracted from the total KB downloaded in step (1).

(4) If applicable, the aggregate KB of annotations are estimated as in steps (2) and (3), and then subtracted from the total.

As a result, we interpret our KB counts as number of KB (and hence, roughly the number of pages) of unannotated statutes, exclusive of headers and footers, but inclusive of tables of contents used to organize those statutes.

---

[22]A typical html page from SD statutes reads:
"*32-5-10.2. Motorcycle safety education fee -- Deposit in special revenue fund. The county treasurer shall remit to the department the motorcycle safety education fees collected pursuant to § 32-5-10.1. The fees shall be deposited in the state treasury in a special revenue fund for use as specified in § 32-20-14.*
Statutes Menu | FAQ | My Legislative Research | Privacy Policy | LRC Menu This page is maintained by the Legislative Research Council. It contains material authorized for publication that is copyrighted by the state of South Dakota. Except as authorized by federal copyright law, no person may print or distribute copyrighted material without the express authorization of the South Dakota Code Commission", where we have italicized the actual statute. The non-italicized portion is 523 bytes, so we subtract 523 bytes per SD www page of law.

[23]The only exception is VT, for which we counted statutes in effect as of 1995. Including year of statute in the regressions has no impact on the results.

Appendix II: U.S. States Data Sources

1920 population by state. University of Virginia Geospatial and Statistical Data Center, plus AK & HI from Texas A&M University (2002).

1990 and 2000 population by state. Census Bureau (2001).

south Alabama, Arkansas, Delaware, Florida, Georgia, Kentucky, Louisiana, Maryland, Mississippi, North Carolina, Oklahoma, South Carolina, Tennessee, Texas, Virginia, West Virginia

lawyers per capita. 1990 Census PUMS weighted number of persons aged 25-54 and reporting working in 1989 and reporting occupation code 178

land area square miles from http://www.imagesoft.net/flags/usstate1.html

90-10 family income differential. 10th and 90th percentiles of the within-state household income distribution, from the March 2001 CPS (referring to year 2000 income). The 90th percentile is divided by the 10th percentile, and the ratio is used in log form in the regressions.

fraction of labor force coded with just 75 occupation codes. Labor force is 1990 Census PUMS persons aged 25-54, reporting work in 1989, and reporting an occupation. The fraction used is the ratio of total labor force persons in a state's 75 largest occupation codes to total labor force.

fraction of labor force employed in agriculture. year 2000 from *Statistical Abstract of the United States* 2001, 2002; Census 2000 Summary File 1, 2 at http://factfinder.census.gov

fraction of population urban, white. year 2000, sources above

income per capita. year 2000, personal income, sources above

number of state senators and representatives. year 2002 from *Book of the States*

year of statehood. http://cointown.com/htm/statehood_facts_2.htm

Appendix Table: Summary Statistics for U.S. State Data
(37 states with KB measures)

variable name	mean	std dev	min	max
KB of law	47,723	21,963	20,922	132,862
2000 population (1000s)	5,970	6,975	493	33,871
1920 population (1000s)	2,165	2,321	55	10,385
south	.27	.45	0	1
lawyers per 1000 people	2.25	0.80	0.70	3.93
area (square miles)	75,632	97,895	1,054	570,833
personal income per cap. (year 2000 $)	28,162	4,491	20,916	40,757
90-10 household inc differential, log	2.07	0.16	1.79	2.42
fraction of labor force coded with just 75 occupation codes	0.70	0.03	0.64	0.77
fraction of population urban	0.72	0.16	0.38	0.94
fraction of population white	0.78	0.14	0.24	0.97
# of state senators & representatives	140	49	49	253
year of statehood	1841	48	1787	1959

Table 1: KB's of Law across States

independent vars	(1)	(2)	(3)	(4)	(5)	addendum: inclusion in sample
log(2000 population)	0.31	0.32	0.36	0.31	0.46	0.34
	(0.06)	(0.07)	(0.09)	(0.02)	(0.06)	(0.24)
south	-0.21	-0.20	-0.22	-0.16	-0.17	-0.37
	(0.10)	(0.09)	(0.10)	(0.06)	(0.07)	(0.23)
lawyers per capita, log	0.20	0.21	0.22	0.16	0.11	-0.05
	(0.10)	(0.10)	(0.10)	(0.07)	(0.11)	(0.24)
land area, log	0.01	0.01	-0.01			-0.13
	(0.04)	(0.04)	(0.05)			(0.12)
inc per cap, log	-0.28	-0.28	-0.23			-0.04
	(0.31)	(0.31)	(0.32)			(0.73)
90-10 household inc differential, log	0.13 (0.22)					
fr. labor force coded with just 75 occ. codes		0.13 (1.53)				
fr. population urban			-0.25			-1.45
			(0.44)			(1.13)
fr of population white	0.19	0.17	0.18			-0.20
	(0.28)	(0.29)	(0.28)			(0.81)
# of state senators & representatives, log	-0.12 (0.11)	-0.12 (0.11)	-0.15 (0.12)			-0.41 (0.24)
year of statehood/100	-0.08	-0.07	-0.05			0.03
	(0.12)	(0.12)	(0.13)			(0.31)
log(1920 population)	0.03	0.03	0.01			-0.16
	(0.05)	(0.05)	(0.06)			(0.15)
adj-R-squared	.82	.82	.82	.84	.85	.13
sample	37	37	37	37	17	50

Notes: (a) dependent variable is log of number of kilobytes of law (c. 2000), except in the last column which is a dprobit for inclusion in our 36 state sample. constant terms are estimated for each specification, but not displayed in the table
(b) specification (5) limits sample to states with year 2000 population at least the median
(c) see Appendix II for variable definitions

Table 2: Northern States, Classified by their Employment Discrimination Laws as of 1968

1950 Black Population	Employment Discrimination Law		
	enforceable	no enforcement commission	no law
avg	171,615	987	492
	100,011	898	375
interquartile range	302,457	1,224	610
avg % of total	1.8	0.2	0.1

enforceable states: AK, AZ, CA, CO, CT, HI, IL, IN, IA, KS, MA, MI, MN, MO, NE, NV, NH, NJ, NM, NY, OH, OR, PA, RI, UT, WA, WI, WY.
no commission states: ID, ME, MT, VT. states w/o a law: ND, SD
Source: Landes (1968)

Table 3: Four Regulations across Countries

dependent variable:	business entry procedures, log	employment laws index	death penalty		have draft	
independent vars	(1)	(2)	(3)	(4)	(5)	(6)
log(population)/10	**0.97** (0.32)	**0.43** (0.29)	**0.50** (0.20)	**0.71** (0.33)	**0.75** (0.24)	**0.20** (0.33)
British legal origin	-0.70 (0.10)	-0.51 (0.09)	0.30 (0.07)	0.40 (0.10)	-0.47 (0.08)	-0.56 (0.10)
real GDP per capita, log	-0.31 (0.08)	-0.11 (0.07)	0.09 (0.04)	0.06 (0.08)	-0.06 (0.05)	-0.02 (0.08)
democracy index	-0.02 (0.18)	-0.01 (0.17)	-0.56 (0.10)	-0.34 (0.19)	0.14 (0.12)	0.01 (0.18)
communist dummy	-0.11 (0.16)	-0.01 (0.14)	0.09 (0.10)	0.39 (0.16)	0.16 (0.11)	-0.03 (0.15)
left power	-0.05 (0.12)	0.10 (0.11)		-0.28 (0.13)		0.16 (0.12)
armed forces per male aged 15-24					1.82 (0.57)	2.04 (0.68)
years at war since 1950					-0.01 (0.02)	-0.01 (0.02)
adj-R-squared	.57	.37	.27	.32	.36	.47
sample	LaPorta	LaPorta	MGX	LaPorta	MGX	LaPorta
countries	71	71	127	71	127	70

Notes: (a) dependent variable is either log of business entry procedures, employment laws index, fraction of years 1976-90 that Amnesty International codes says there was a "retentionist" death penalty policy, or fraction of years 1985, 90, 95 having conscription.
(b) OLS standard errors are in paren.
(c) democracy index is on 0-1 scale, and averaged 1975-90

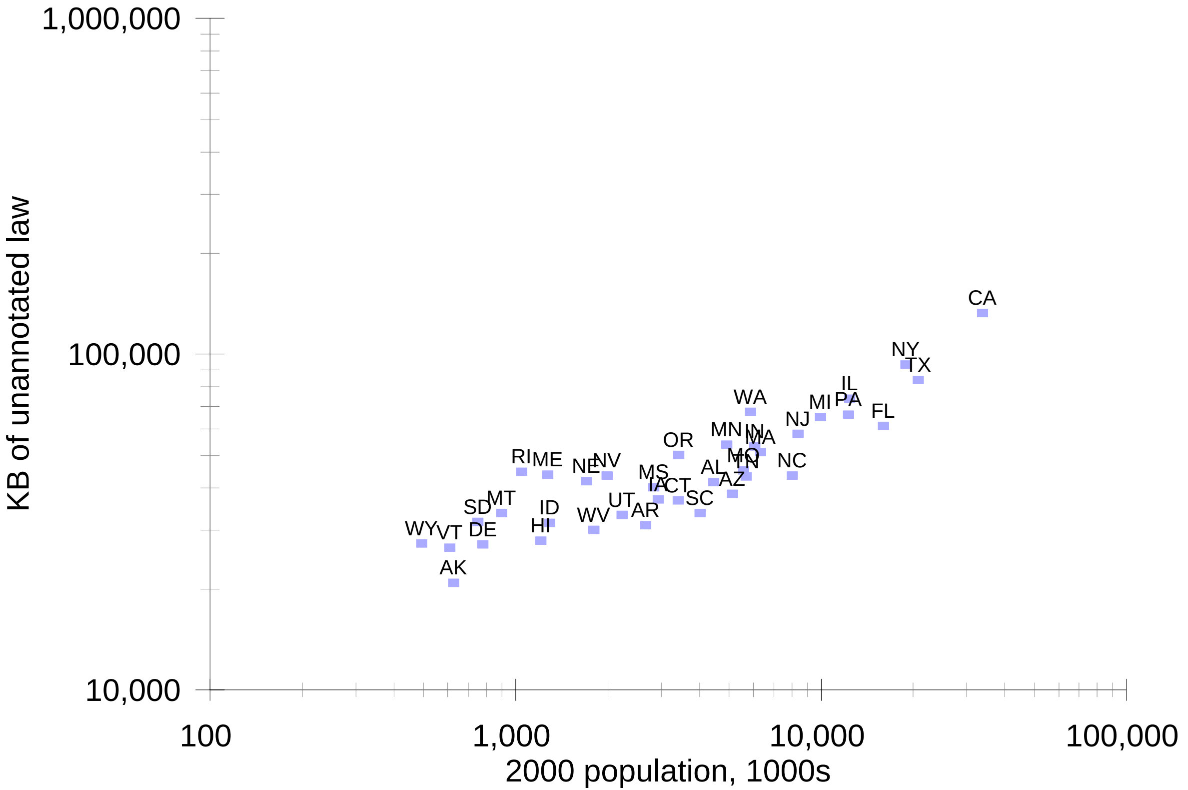

Figure 1. Statutes and Population Across States

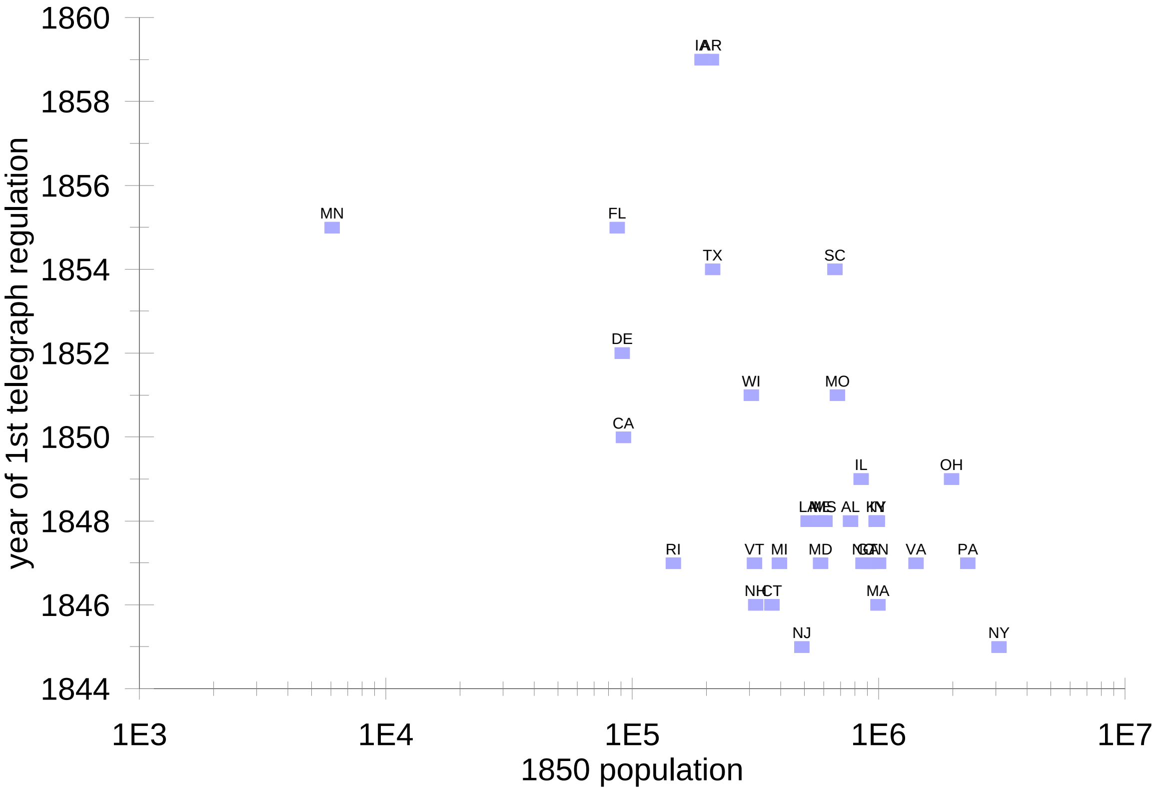

Figure 2 Telegraph Regulation Across States

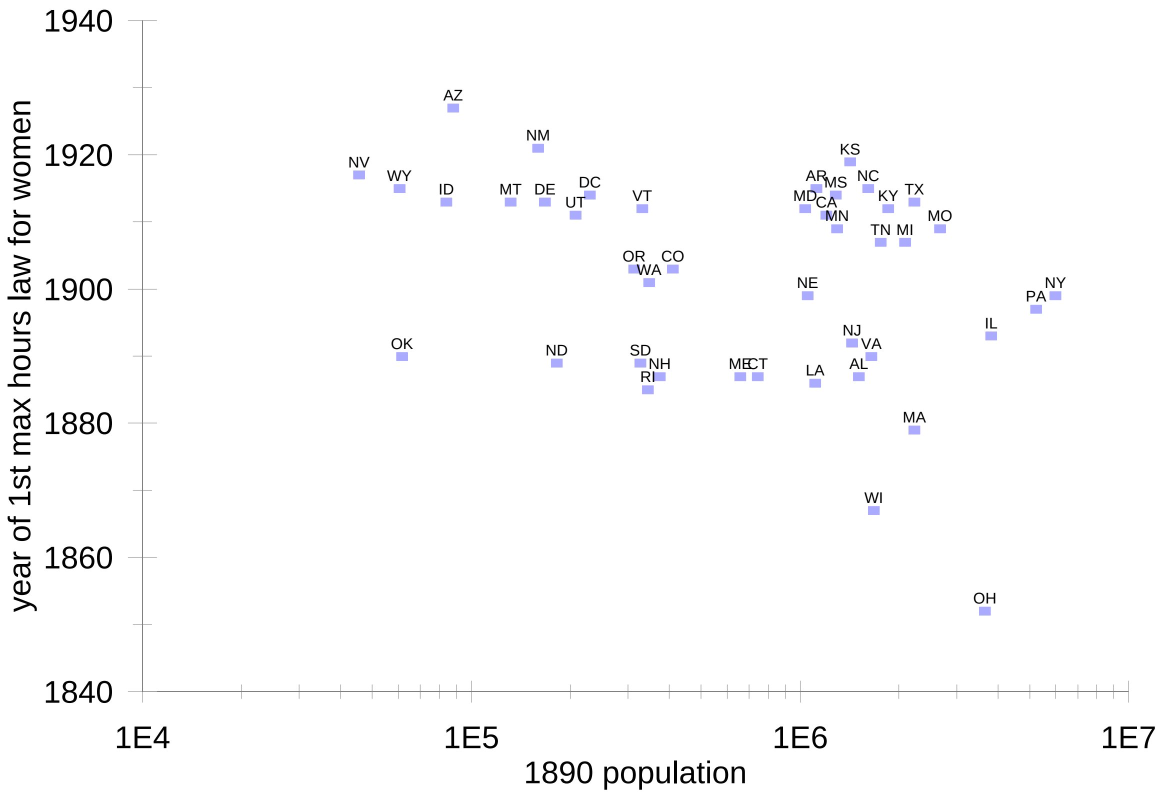

Figure 3 Female Hours Laws vs Population

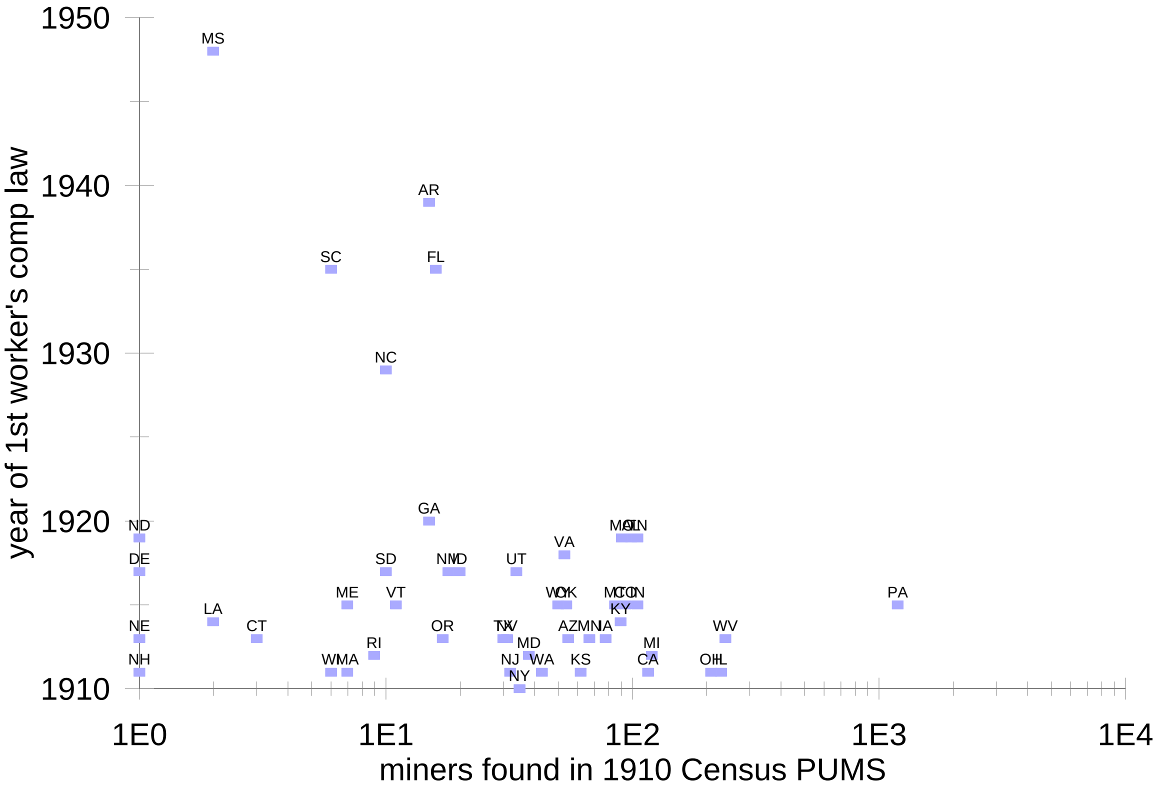

Figure 4 Total Miner Population and the Adoption of Workers' Comp

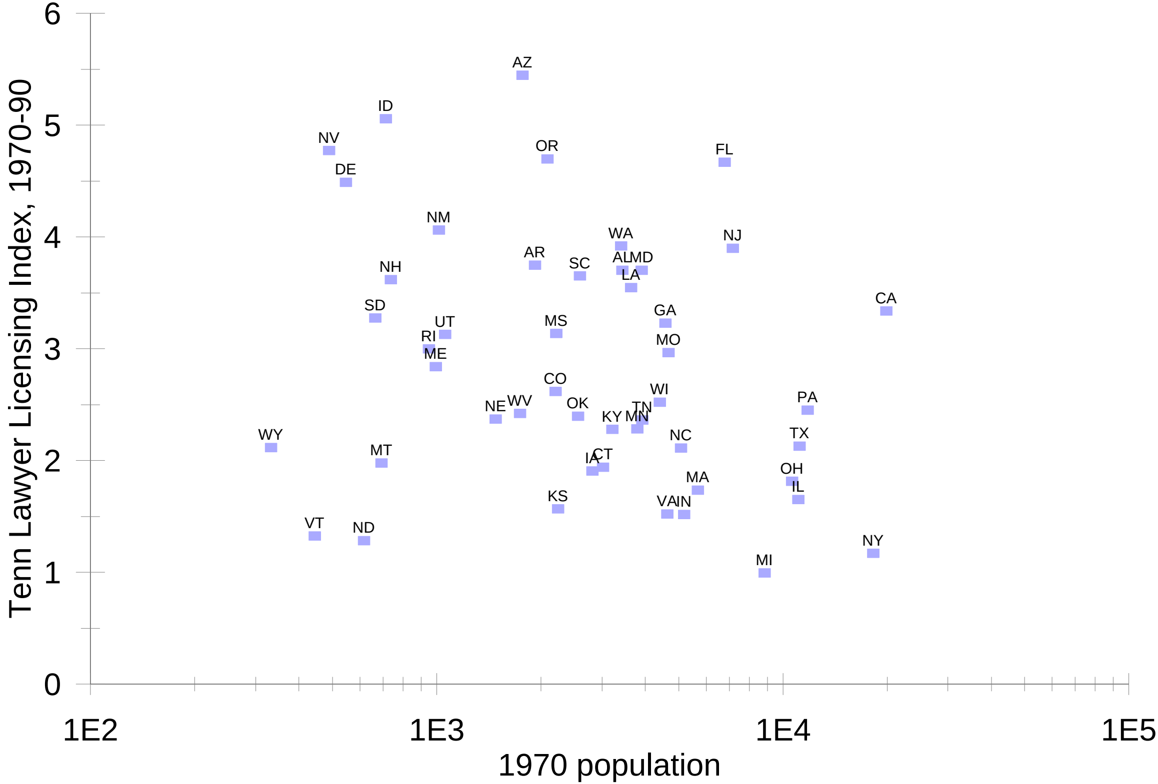

Figure 6 Population and the Strictness of Lawyer Licensing