Legal centralization and the birth of the secular state
Noel D. Johnson, Mark Koyama*

Center for the Study of Public Choice, George Mason University, 4400 University Drive, Fairfax VA, United States

A R T I C L E   I N F O

Article history:
Received 28 December 2012
Revised 29 July 2013
Available online xxxx

JEL classification:
H10
N43
Z12
K42
P48

Keywords:
State capacity
Religion
Secularization
Heresy
Legal capacity

A B S T R A C T

Johnson, Noel D., and Koyama, Mark—Legal centralization and the birth of the secular state

This paper investigates the relationship between the historical process of legal centralization and increased religious toleration by the state. We develop a model based on the mathematics of mixture distributions which delineates the conditions under which legal centralization raises the costs faced by states of setting a narrow standard of orthodox belief. We compare the results of the model with historical evidence drawn from two important cases in which religious diversity and state centralization collided in France: the Albigensian crusades of the thirteenth century and the rise of Protestant belief in the sixteenth century. Journal of Comparative Economics xxx (xx) (2013) xxx–xxx. Center for the Study of Public Choice, George Mason University, 4400 University Drive, Fairfax VA, United States.
© 2013 Association for Comparative Economic Studies Published by Elsevier Inc. All rights reserved.

1. Introduction

‘Religious freedom is not just one liberal freedom among others . . . it is the prototypical liberal freedom, a cornerstone of modern political rights’ (Blackford, 2012, 1). But how did religious freedom and a commitment to religious toleration emerge? In premodern Europe, states legitimized their authority by imposing limits on acceptable religious beliefs. In contrast, governments in the developed world today – even in countries where there is an official state church or where levels of religiosity are high – tolerate a wide range of beliefs and practices when compared with governments of the past. Conventionally, the separation of church and state is dated to the ratification of the First Amendment in the United States in 1791. But the modern origins of religious toleration and freedom go back to the period of the Reformation in Europe.1 This paper provides a theoretical mechanism that helps to explain this transition from repression to toleration and from the religiously sanctioned polities of the past to the modern secular state.

We emphasize the role played by state formation and legal centralization in leading European states to gradually widen the range of permissible religious beliefs during the period following the Reformation.2 We develop a simple

* Corresponding author.
E-mail addresses: njohnsoL@gmu.edu (N.D. Johnson), mkoyama2@gmu.edu (M. Koyama).


2 As such, we contribute to the growing literature on the importance of state capacity for modern economic development (Besley and Persson, 2009, 2010, 2011; Acemoglu et al., 2005; Johnson and Koyama, 2011, 2013).
decision-theoretic model which allows rulers to invest in both fiscal capacity and in maintaining religious orthodoxy in order to increase their legitimacy. It predicts that rulers facing populations with more diverse religious beliefs are less likely to rely on religious legitimation in order to lower the cost of collecting taxes. The model also shows how this effect depends on the extent to which fiscal capacity and religious legitimacy are substitutes. We use our framework to explore what happens when a common legal or religious system is imposed on previously independent regions with different underlying religious beliefs. The model predicts that when religious beliefs in the two regions are sufficiently salient and diverse then the process of legal standardization will give rise to greater de jure religious toleration.

We provide contemporary cross-country evidence that is consistent with our model. But since these data do not shed light on the mechanisms responsible for generating the correlations we observe, the rest of the paper employs an analytic narrative approach (as in, for instance, Greif, 2006). This narrative focuses on two detailed case studies from French history to provide support for the model’s predictions. We also evaluate our claims in the context of historical evidence from medieval and early modern Europe more generally.

Our first detailed case study shows how increased state centralization can lead to greater religious repression. We show why the thirteenth century incorporation of Languedoc into France led to local beliefs that had previously been tolerated becoming designated as heretical. The second case study discusses the rise of Protestantism in sixteenth century France. We argue that it was the conjunction of two factors: the commitment of the French state to a standardized legal system and the emergence of Protestantism as particularly powerful and cohesive body of religious deviants that made a policy of religious persecution extremely costly. As a result of the high costs of religious persecution, Protestantism was decriminalized in 1563 and a few decades later the Edict of Nantes (1598) was promulgated acknowledging the co-existence of two religions within a single state. Thus, investments in state capacity were accompanied by increased religious toleration in France in the sixteenth century, not because individuals were more enlightened, but simply because enforcing an intolerant state orthodoxy was too costly. We provide evidence that similar processes were at work in Germany and in England where the Peace of Augsburg (1555) and Act of Uniformity (1559) respectively marked an end to large-scale heresy trials and served as examples which shaped the gradual emergence of religious toleration in the West during the seventeenth and eighteenth centuries.

Our arguments build on recent work which sheds light on the conditions under which minority groups (religious or ethnic) are likely to face persecution. Gill (2008) focuses on the contrasting experiences of North and South America in the colonial period. He argues that because North America had a greater diversity of religious beliefs than did South America, a higher degree of religious freedom developed there as persecution was prohibitively costly. Jha (2013) examines why trading cities in southern India were able to sustain religious toleration for Muslim minorities. He argues that religious toleration is more likely to emerge if a minority group provides non-replicable and non-expropriable economic services. Minorities who lack such a niche or whose economic niche is more easily substitutable are more likely to face persecution. Anderson et al. (2013) study the relationship between negative economic shocks, as measured by extreme temperature variations, and the expulsion and persecution of Jews in medieval and early modern Europe. They suggest that the emergence of stronger states in the seventeenth century provides one explanation for the decline in the number of antisemitic persecutions.

There are several alternative hypotheses that seek to explain the rise of religious toleration. One influential view attributes the growth of toleration to the Enlightenment of the eighteenth century. In contrast, we argue that important movements towards toleration occurred prior to the writings of leading Enlightenment thinkers such as Locke and Voltaire, at a time when very few writers advocated tolerance. Another explanation relies on the doux commerce hypothesis which associates the rise of toleration with the growth of markets and trade (Hirschman, 1977). While we believe this is important in explaining the gradual development of more tolerant attitudes at an individual level, it does not explain the relatively rapid changes in policy by major European states in the sixteenth and seventeenth centuries—a time when popular religious hatreds remained strong.

There are also large literatures in sociology, political science, and economics evaluating the secularization hypothesis. However these tend to focus on the demand for and supply of religious goods and not on the rise of religious toleration which is the focus of this paper. The most relevant papers in the economics of religion literature are Barro and McCleary (2005) and Cosgel and Miceli (2009a). Using cross-country panel data from the nineteenth and twentieth centuries, Barro and McCleary (2005) examine why some countries have state churches while others do not. They find that state religions are established where there is a high degree of religious homogeneity and suggest that countries with more heterogeneous religious preferences are more likely to have greater separation between church and state. In this paper we seek to provide an analytical mechanism that can explain this correlation. In so doing, we build on the argument that religious legitimation played a crucial role in enabling premodern rulers to obtain the compliance of the population (Cosgel and Miceli, 2009a; Cosgel et al., 2012). What is new in our theory is that we explain under what conditions the process of state-building and legal centralization generates increased religious toleration.

More generally, our analysis relates to the literature on the optimal size of nations (Alesina and Spolaore, 1997; Alesina and La Ferrara, 2005a,b; Alesina and Reich, 2013). Whereas this literature focuses on the how political borders adjust to accommodate diverse preferences, however, we study how a ruler’s attitude to enforcing religious or ideological

---

3 Important sociological works on this topic include Bruce (2002) and Norris and Inglehart (2004). The supply-side argument against the secularization hypothesis is discussed by Iannaccone et al. (1997).

4 Cosgel and Miceli (2009a) build a model of the relationship between a state and religion. In their model autocratic states have an incentive to support religions that lend their regime legitimacy and thereby lower the cost of obtaining tax revenue. Also see Rubin (2011).
uniformity varies under different distributions of religious beliefs. Our model highlights this process and outlines the relevant costs and benefits that determine the state’s decision to either forcibly convert the dissenting population or expand the bounds of religious toleration.

2. State capacity, religious fragmentation, and tolerance today

To motivate our thesis, we first look at the empirical relationship between religious diversity and religious freedom today. We combine data on cross country religious fractionalization from Alesina et al. (2003) with data on government regulation of religion, fiscal capacity, and other controls from the Cross-National Socio-Economic and Religion Data (2011). The religious fractionalization data measure the probability that two randomly drawn individuals from a population will hold different religious beliefs. The government regulation of religion variable is an index that covers six aspects of government’s involvement in religion such as interference with the right to worship, the regulation of mission work, proselytizing, preaching and conversion, and the presence of legal codes and policy actions related to religious belief (Grim and Roger, 2006). We split the resulting sample of 129 countries into low fiscal capacity and high fiscal capacity governments depending on whether they are above or below the mean in taxes collected per capita. After controlling for income, education, and institutional quality, we then look at the non-parametric relationship between religious fractionalization and government regulation of religion in Figs. 1 and 2.

These data are suggestive but they do not permit a clean empirical test of the causal relationship between state building and religious toleration. For this reason, in what follows we develop a simple model and historical narrative that elucidates the relationship between state capacity, religious fractionalization, and religious freedom. Our model will allow us to generate testable predictions about the conditions under which legal centralization will either lead to greater or lesser state toleration of religious beliefs. Together with our historical evidence and case-studies we will then have a fuller understanding of the relationships presented in Figs. 1 and 2.

3. A model of religious legitimacy and toleration

In this section we develop a decision-theoretic model to elucidate the incentives a ruler has to increase or decrease religious toleration. The framework parallels the literature on the optimal size of nations in exploring how heterogenous preferences determine whether or not regions should integrate (Alesina et al., 2004; Alesina and La Ferrara, 2005a; Spolaore, 2009). What is novel about our approach is that instead of focusing on the endogenous creation of political borders, we focus...
on the determination of the ideological borders of tolerance and take explicit account of the way the distribution of beliefs can change when jurisdictions are unified.

We begin with $n$ states each governed by an autocratic ruler indexed by $j$. Each sovereign rules over a large population of individuals indexed by $i$. Individual $i$'s religious belief is characterized by $b_{ij} \in \mathbb{R}^2$. This measure captures the ideological distance of a person from some baseline standard of orthodox belief. Thus, for example, it could measure adherence to orthodox traditions (e.g. eating certain foods, wearing certain clothes, etc.) or whether an individual believes in the dogma of the orthodox church (e.g. belief in consubstantiation or transubstantiation). The belief of individual $i$ in region $j$ is drawn from probability density function $B_j(b) \sim \mathcal{N}(\mu_j, \sigma_j)$.

3.0.1. The Ruler's Maximization Problem

We assume that each ruler $J$ is revenue-maximizing, as in Brennan and Buchanan (1980). He obtains tax revenue $R_j$ which is determined by the optimal level of taxation $T_j$ (assuming all subjects comply) and the level of non-compliance or evasion $F_j$:

$$R_j = \left(1 - F_j\right)T_j,$$

where $T_j$ is the optimal level of taxation in region $j$.

The level of non-compliance is a function of two factors. Following the literature on state capacity, we argue that one determinant is the level of fiscal capacity $q_j \in \mathbb{R}^+$.\(^8\) Similar to Cosgel and Miceli (2009b) and Cosgel et al. (2012) we argue that the ability of the state to collect taxes also depends on its religious legitimacy $(\lambda_j \in \mathbb{R}^+)$ in what follows we assume that these factors determine compliance according to the following CES production function:

$$1 - F_j = [\alpha(q_j)^{\beta} + (1 - \alpha)(\lambda_j)^{\beta}]^{1/\beta},$$

where $\alpha \in [0,1]$ is the share parameter and $\beta < 1$ determines the degree of substitutability between fiscal capacity and religious legitimacy. Total revenue for ruler $j$ is therefore equal to:

$$R_j = T_j\left([\alpha(q_j)^{\beta} + (1 - \alpha)(\lambda_j)^{\beta}]^{1/\beta}\right).$$

A ruler can increase revenue for a given rate of taxation by investing in greater fiscal capacity (higher $q$) or gaining greater religious legitimacy (higher $\lambda$). Both of these policies involve costs, however. Investment in fiscal capacity has a marginal cost of $\kappa$. Reflecting the fact that increased reliance on religious legitimacy results in increased persecution of individuals with unorthodox beliefs, we define its cost in terms of the number of people persecuted (defined as heretical) in society $j$. This relationship will be more precisely defined in the next section.

3.1. Religious toleration

We focus on the ability of the ruler to increase religious legitimacy by setting the officially tolerated bounds of belief narrowly, thereby signalling his or her commitment to the religion. We assume officially tolerated beliefs lie on an interval $[\mu_j - \sigma_j, \mu_j + \sigma_j]$.

---

\(^8\) $\rho$ is most easily thought of as the monitoring and collection apparatus available to the ruler. For a detailed discussion see Johnson and Koyama (2013) and the citations therein.
\( \tau \in \mathbb{R}^+ \) constructed around the ruler’s definition of canonical belief, which is located at the mean of the distribution. We define religious legitimacy as \( \lambda = 1/\tau \). The more narrowly the bounds of toleration are defined, the more legitimacy religion brings to the tax apparatus of the state.\(^9\)

We define the proportion of the population defined as heretical as \( A \in [0,1] \). Under our assumption that the distribution of belief \( B_j \) is normal with standard deviation \( \sigma_j \), the value of \( A \) conditional on \( \lambda \) is defined by,

\[
A = 1 - \text{erf} \left( \frac{1}{\sqrt{2\lambda \sigma_j}} \right), \tag{4}
\]

where \( \text{erf} (\cdot) \) is the error function.\(^{10}\) Since the first derivative of the inverse error function is positive, there is a positive relationship between \( \lambda \) and \( A \). Furthermore, ceteris paribus, an increase in \( \sigma \), the dispersion of religious belief, also results in an increase in \( A \).

We assume that persecuting large numbers of people becomes more costly the more narrowly the ruler defines the bounds of tolerance. Thus, we write the total cost of investing in religious legitimacy as \( k \cdot A(\lambda, \sigma) \) where the price of investment in legitimacy, \( A(\lambda, \sigma) \), is simply the proportion of the population that is being persecuted. We can then write the budget constraint of a ruler as follows:

\[
k \cdot p + \lambda \cdot A(\sigma, \lambda) = R,
\]

where \( R \) is the amount of the budget allocated to investing in enforcement of tax collection (either increasing \( q \) or \( k \)). This budget constraint is required to solve the ruler’s maximization problem but we do not appeal to shifts in \( R \) in our analysis as it is an exogenous parameter.

This setup emphasizes several points. First, the more religiously diverse a society is (\( \sigma \) larger), the more costly is reliance on religious legitimacy. Second, the degree of persecution of deviant religious beliefs in a society (the size of \( A \)) is distinct from a society’s degree of tolerance (the size of \( \lambda \)). It is possible for two regions to persecute the same proportion of their population, but the region with a greater variance of belief will, in equilibrium, be more tolerant. Fig. 3 illustrates this result.

In the Figure, Region 2 is more tolerant than Region 1 in the sense that it defines orthodox belief more broadly than does Region 1. However, in equilibrium, both regions are persecuting the same proportion of their populations as heretical. In effect, the government in Region 2 is forced to accommodate larger deviations from the orthodox belief because it has a more heterogenous population.

### 3.2. Equilibrium and comparative statics

Each ruler \( j \) wants to maximize total revenue subject to the costs of investing in fiscal capacity and maintaining religious legitimacy. In equilibrium, the revenue maximizing ruler will invest such that the marginal rate of technical substitution between fiscal capacity and religious legitimacy is proportional to the ratio of their costs. This is described by the following condition:

\[
\frac{\rho}{\lambda} = \left( \frac{1 - \kappa}{\kappa} \cdot \frac{\kappa}{A + \lambda \cdot A_j} \right)^{\frac{1}{1+\kappa}}, \tag{5}
\]

where \( A_j \) is the partial of \( A \) with respect to \( \lambda \). Eq. (5) has a fairly straightforward interpretation. As the price of investment in fiscal capacity (\( \kappa \)) increases, the ruler shifts investment towards religious legitimacy. The term \( (A + \lambda \cdot A_j) \) represents the

---

\(^9\) Implicitly we are assuming that the ruler of a territory can control the extent to which religious beliefs are tolerated. We should note that in medieval Europe, heresy was the responsibility of religious and not secular courts. The decision to initiate a campaign of persecution against heretics was, however, almost always made with the consent of the secular powers (Moore, 1978).

\(^{10}\) The error function is defined as: \( \text{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2} \, dt \).
marginal cost of investing in legitimacy and is strictly increasing in the variance of belief (\(\sigma\)).\(^{11}\) Thus, as \(\sigma\) increases, the ruler will substitute away from reliance on religious legitimation in favor of investing in fiscal capacity.\(^{12}\)

Eq. (5) has implications for equilibrium religious toleration (\(\tau^*\)) which we summarize in the following proposition.

**Proposition 1. Equilibrium religious toleration \(\tau\)**

(i) is increasing in the variance of religious belief \(\sigma^2\);

(ii) is decreasing in the cost of building fiscal capacity \(\kappa\);

(iii) as the degree of substitutability between fiscal and religious capacity \(\beta\) increases, an increase in the dispersion of religious beliefs \(\sigma\) leads to a proportionally larger increase in religious toleration \(\tau\).

The first two parts of Proposition 1 follow directly from the maximization problem. The third part of Proposition 3 is more subtle and has less obvious implications. It states that if fiscal capacity and religious legitimacy are highly substitutable sources of political authority then small increases in the variance of religious beliefs can generate large increases in religious toleration. We treat \(\beta\) as a parameter in our model but in our historical application it makes sense to think of it as slow moving quasi-parameter in the sense of Greif and Laitin (2004). In the middle ages there were a few alternative sources of political legitimacy. But, as Greif and Rubin (2013) argue, in early modern Europe alternative sources of political legitimacy became more prominent (in England, for instance Parliament took on this role in the sixteenth and seventeenth centuries). It is possible to think of these developments as increasing \(\beta\) and thereby making an increase in the variance of religious belief, \(\sigma\), more likely to lead to a proportionately larger shift towards religious toleration than would otherwise have been the case.

3.3. Legal centralization across different jurisdictions

We now use our model to examine what happens to toleration when one of the regions, which we call the central region, standardizes laws across the other regions. In particular, we wish to investigate the difference between religious tolerance when each local government sets its own legal rules and the level of religious tolerance that results from a single centralized government imposing legal standardization.

We do not explicitly model the decision of a ruler to impose legal centralization as a number of recent papers explore this trade-off. Gennaioli and Voth (2011), for example, develop a formal model to examine the relationship between military conflict and the decision to centralize fiscal and legal institutions in early modern Europe. We build on this work by taking the decision to centralize as given, and then studying its consequences for religious toleration.

For clarity, we refer to variables associated with the centralizing region before centralization with the letter \(c\). After centralization, we index variables associated with the single unified government with the letter \(m\). As above, the central government chooses a range of legitimate belief centered on the orthodox standard, which is defined as the set of beliefs closest to that of the central mass of the population. We assume, without loss of generality, that the underlying costs and benefits of persecuting a certain proportion of the population are the same for the centralized government as they are for the regional governments. The main assumption that makes the centralized government different from the regional government is that its definition of tolerance, \(\tau_m\), must be imposed on all regions. This means the central authority is choosing \(\tau_m\) for the distribution of the distributions of belief across all the regions. This distribution is called a ‘mixture distribution’ and is defined as the convex combination of the underlying distributions. We write this as

\[
M(b; \mu_1, \ldots, \mu_n, \sigma_1, \ldots, \sigma_n) = \sum_{j=1}^{n} \omega_j \beta_j(b; \mu_j, \sigma_j),
\]

\(11\) To see this, note that the partial of \(A + \lambda A_1\) with respect to \(\sigma\) is equal to:

\[
\frac{\partial}{\partial \sigma} \left(1 - \text{erf} \left(\frac{1}{\sqrt{2}}\right) + \lambda \left(1 - \text{erf} \left(\frac{1}{\sqrt{2}}\right)\right)\right) = \sqrt{\frac{2}{\pi}} \text{erf} \left(\frac{1}{\sqrt{2}}\right) \frac{2}{\sqrt{\pi}}. \text{The term on the RHS is always positive for positive values of } \lambda \text{ and } \sigma.
\]

\(12\) As \(\beta \to 0\) fiscal capacity and religious legitimation become perfect substitutions, while as \(\beta \to 1\) religious legitimation and fiscal capacity become perfect complements; if \(\beta \to -\infty\) Eq. (2) becomes Cobb Douglas.
where $w_j$ is a weight assigned to each region such that the sum of all $n$ weights is one. A natural interpretation of these weights is as the relative importance of a region in the overall distribution (e.g. population share, proximity to the center, etc.). Fig. 4 shows an example of a mixture distribution along with its two underlying distributions.

The cumulative distribution function of $M$ also exists and is constructed as a convex combination of the underlying CDF's. Since $M$ is a mixture of normal distributions, furthermore, it is established in the literature that the inverse of its CDF also exists (Everitt and Hand, 1981). Thus, there is a unique mapping of $A$ to $s_m$ similar to Eq. (4). Unfortunately, there are no closed form solutions for the inverse CDF of a mixture of Gaussian distributions and, thus, it is not possible to write this down as an explicit expression for $s_m$. There are closed form expressions for the mean and variance of $M$, however. The mean is simply the weighted sum of the means of the component regions:

$$\mu_m = \sum_{j=1}^{n} \omega_j \mu_j.$$  (7)

Similarly, the variance of the mixture distribution is the weighted sum of the underlying regional variances with an adjustment for how much each region is ‘off-set’ from the others:

$$\sigma^2_m = \sum_{j=1}^{n} \omega_j \left[ (\mu_j - \mu_m)^2 + \sigma_j^2 \right].$$  (8)

Eq. (8) makes clear the potential for fat tails in the mixture distribution. We can also use Eq. (8) to draw some conclusions about the effect of legal centralization on religious tolerance.

For concreteness, we define an increase in religious toleration as occurring after legal centralization if $s_m \geq \max \{ s_j \}$. Furthermore, we say region $j$ has more ‘cohesive’ religious beliefs than region $k$ if $\sigma_j < \sigma_k$. We say region $j$ has more ‘deviant’ religious beliefs than region $k$ if $(\mu_j - \mu_m)^2 > (\mu_k - \mu_m)^2$. We can then state the following based on the existence of $s_m$ and Eq. (8).

**Proposition 2.** An increase in religious toleration is more likely to accompany legal centralization the more cohesive and deviant are the religious beliefs of the underlying regions.

As there exists no closed form solution for a mixture of Gaussian distributions we are unable to provide an analytical proof of this Proposition. Instead we establish it numerically and graphically.

To demonstrate Proposition 2 numerically we use the statistical package R to calculate the quantile function for the mixture distribution of two regions. We hold the distribution of beliefs constant for the central region at $\mu_c = 0$ and $\sigma_c = 1$. For the other region we allow $\mu_j$ to vary between 0 and 100 and $\sigma_j$ to vary between 1 and 100. In Fig. 5 we plot the different combinations of cohesiveness and deviance for region $j$ which result in $s_m = s_j$. If we assume $A$ is equal to 0.05, for example, all

---

The potential for large higher-order moments (e.g. skewness, multiple modes, etc.) is well known in the literature on mixture distributions (Ray and Lindsay, 2005; Titterington et al., 1985). Indeed, the ability of mixtures to accommodate high degrees of skewness and kurtosis is one of the main reasons they are frequently employed in the field of empirical finance (Tsay, 2005).

This parameterization facilitates interpretation since tolerance in region $j$ is always weakly greater than tolerance in the central region.
combinations of cohesiveness and deviance for region $j$ that lie above the plotted line result in a mixture distribution which yields increases in toleration after legal centralization. Likewise, combinations of cohesiveness and deviance which fall below the line will result in a decrease in toleration after centralization. Consistent with Proposition 2 there is a positive relationship between both cohesiveness and deviance and the likelihood of increased toleration after legal centralization. If both regions have the same cohesiveness of religious belief (equal variances), then any difference in their means will result in increased toleration after centralization. Furthermore, as the cohesiveness of a region decreases, toleration becomes less likely. This relationship is exacerbated for higher values of $A^*$. In other words, the more religious persecution there is initially, the less likely it is that legal centralization will lead to greater secularization.

We illustrate Proposition 2 graphically in Figs. 6 and 7. The distribution of belief in the central region is the same in both figures. However, in Fig. 6 region $j$ has both low cohesiveness and low deviance. In effect, there is only a weak consensus among believers concerning what is orthodox in $j$, but at the same time, the average believer holds views consistent with those of the central region. As a result, when the central region imposes a common legal standard, the bounds of tolerance are reduced in region $j$. This is a case in which legal centralization is accompanied by a decrease in secularization.

In Fig. 7, region $j$ has both high cohesiveness and high deviance. Consequently, when legal centralization occurs, secularization increases. This is true since the new standard of tolerance, $r_m$, is greater than the previous value of tolerance in the most tolerant region before centralization which, in this case, happened to be the centralizing region.

Finally, while we have developed a static model to study the incentives centralizing rulers face with regards to religious diversity, it is possible to think about how the insights of this framework would generalize to a dynamic setting. In particular, even in cases where our static approach predicts that an increase in the variance of beliefs $\sigma$ will give rise to greater tolerance in equilibrium, out of equilibrium, there may be a period of greater persecution. Periods of intense state-building are often accompanied by spikes in persecutions for heretical beliefs. Our model is consistent with this behavior if we assume that the centralizing region initially imposes its own standard of tolerance on the unified regions and if the central authority is slow to adjust this standard after heresy trials increase.
This slowness can be thought of in terms of adjustment or information costs. If this is the case then the process of centralizing will increase the number of persecutions, regardless of whether equilibrium toleration has increased or decreased so long as $\tau_c < \tau_m$. Fig. 8 illustrates the logic behind this claim. The initial standard of tolerance after unification is that of the centralizing region, $s_c$. After unification, this standard results in a number of prosecutions in excess of what is optimal from the point of view of the government. If the central authority only adjusts $\tau$ with a lag, then depending on how significant this lag is, heresy prosecutions could exceed what is optimal for an extended period of time. Eventually, however, the logic of Eq. 5 will put pressure on the government to decriminalize a significant amount of heretical beliefs by expanding tolerance to its equilibrium value of $\tau_m$.

4. The theory applied to history

In this section we apply our theoretical model to two historical case studies: the suppression of the Cathars in southern France in the thirteenth century and the failed attempt to suppress Protestantism in France in the sixteenth century. While they do not rule out other causal factors, these case studies provide historical evidence that the mechanism identified in our model played an important role in explaining the rise of religious toleration that took place in Europe in the sixteenth and seventeenth centuries.

4.1. The creation and suppression of medieval heresies

In medieval Europe, the Catholic Church played a crucial role in legitimating secular authority. In theory, therefore, religious toleration was impossible as it meant undermining the Church’s monopoly over religious authority. The motto of the French king: un roi, une loi, une foi applied across Christendom. Nevertheless, the ability of either medieval church or state to enforce religious orthodoxy in practice was limited. Robert Moore observes the ‘structure of the Western Church itself in the early middle ages was one which permitted, and was bound to permit, a much greater variety than would later be thought consistent with the maintenance of Catholic unity. It had not yet developed the means, or, some would say the inclination, to demand uniformity of worship and practice throughout Western Christendom’ (Moore, 1978, 65).

Our model emphasizes the distinction between the bounds of religious toleration within a society ($\tau$) and the extent to which that society persecutes deviant belief ($A$). This distinction is necessary to understanding the relative absence of religious persecution in the early middle ages. Secular rulers in the early middle ages were not necessarily tolerant, but they lacked the capacity to impose a strict standard of belief. As a result, large-scale, state-sanctioned, judicial killings of individuals for their beliefs were rare. After 1150 this began to change; monarchs ‘equipped themselves with institutions of governance that were increasingly professionalized in their staffs and bureaucratized in their operations … jurists and learned writers bent their efforts to the rationalization of legal systems’ (Given, 1989, 336).

At the same time, rulers became increasingly reliant on religious legitimacy to justify their rule. In the language of our model this can be thought of as an exogenous downwards shift in the coefficient of substitution between fiscal capacity and religious legitimation $\beta$. We do not attempt to explain this development as it not fully understood by historians and

---

15 In practice, some toleration could be permitted through the use of exemptions, for example for Jews (see discussion in Koyama, 2010), and for Muslims in lands reconquered by crusaders. Religious justification for the use of force in coercive belief was provided by St. Augustine (see Atkins and Dodaro, 2001).

16 See also Wakefield (1974, 15), and Lambert (1998, 17). Priscillian of Avilia was sentenced as a heretic, or witch, in 383. No one was again executed for heresy until 1022 and then no heresy trials resulted in executions between 1022 and 1143 (Moore, 1978, 13–23). Note, however, that heretics were 'burnt by the people' in Cambrai in 1077, Soissons in 1114, and Liège in 1135 (Moore, 2012, 7).

17 Joseph Strayer observes: 'It is not surprising that the regnum had little resemblance to a state, for, in the early middle ages, it is doubtful that anyone had a concept of a state.' (Strayer, 1963, 1971b, 342).
was driven by slow moving cultural factors. Medieval rulers came to use religion to justify their authority and made themselves responsible for campaigning against heresy. Moore (1978) has influentially termed this the birth of the persecuting society.9

Nowhere was this tendency more marked than in France; the French king was consecrated as well as crowned on his coronation ceremony, and as part of his coronation oath he swore 'that in good faith to all men I shall be diligent to expel from my land and also from the jurisdiction subject to me all heretics designated by the Church' (Holt, 2005, 7). From the reign of Philip Augustus (1179–1223) onwards, he cultivated the title rex Christianissimus (Most Christian Monarch) and claimed that his land was the freest from heresy in all of Christendom. (see Strayer, 1969, 1971a, 308).20

Furthermore, the process of state building in medieval Europe led to the integration of fragmented regions, with different cultures and different religious traditions and customs into large proto-nation states such as France and England. This was particularly marked in the case of France, where the south, Provençal or Languedoc, was linguistically and culturally distinct from the north.21 The institutions of the Catholic Church were less developed in Languedoc than in northern France.22 This disparity between local religious practices and Catholic orthodoxy and the failure of local rulers to suppress these ‘heretics’ became the justification for the Albigensian Crusade (1209–1229), the largest campaign against heresy in medieval Europe.

The heretics suppressed in thirteenth century Languedoc are usually labelled Cathars. Traditional historical accounts view the Cathar heresy as a distinctive dualist religion influenced by eastern heresies and by older Gnostic and Manichean ideas.23 However, the most recent historical scholarship casts doubt on the existence of Catharism as a distinctive set of beliefs prior to the crusade and the introduction of the Inquisitorial method. The best evidence we have on the beliefs of ordinary people comes from the Inquisitorial records. The largest Inquisitorial investigation of heresy of the medieval period was conducted in Languedoc in 1244–1245 and revealed no trace of an alternative, organized, dualistic religion (Pegg, 2001). Instead, what this research suggests is that it was the very process of state formation and legal centralization that led to beliefs that were previously considered within the limits of accepted Catholicism becoming defined as heretical.

The heretics were not called Cathars or Albigensians in the contemporary literature, instead they were often referred to as the good men or bons omes (see Pegg (2001, 19) and Moore (2012, 302)). Importantly, ordinary people, even those who declared themselves heretics, did not necessarily have a clear idea of what ideas distinguished the ‘good men’ from the Catholic Church.25 Pegg concludes that ‘no systematic dualist philosophy comes through the confessed thoughts of all these men and women; rather, instead of valid generalizations, a collection of vague generalities appear to have been recorded at Saint-Sernin’ (Pegg, 2001, 79). Prior to the crusade against them, the heretics of the Languedoc did not necessarily reject the Catholic faith: they were often buried with orthodox Catholics and participated in the Catholic sacrament.

9 See Moore (2012) for a discussion of the centralization of the Church in the eleventh century as a result of the Gregorian Reforms. These reforms enhanced the power and authority of the Church.

10 Henry II was an early secular ruler to take heretical belief seriously by, for example promulgating the Assize of Clarendon in 1166. The Holy Roman Emperor Frederick II was the first secular ruler to legislate that all heretics should be burnt in 1224; other rulers followed him. Louis VIII of France enacted anti-heresy legislation in 1226 barring heretics from holding public office and appropriating their lands. In 1233 anti-heresy legislation was incorporated into the laws of Aragon (Hamilton, 1982).

11 Walter Ullman observes that the ‘acquittal of the theocratic functions in the French king could hardly be better demonstrated to contemporaries and also illustrated to later generations than through the elaborate coronation ceremonial. In fact, it would be difficult to point to a coronation symbolism which was better thought out, more elaborate and indicative, than the French ordines. Each succeeding recension of the ordre improved its predecessors by symbolic details and expansion of individual features. Here we can only briefly refer to the one or the other element. The holy oil of Clovis – the sainte ampoule – was one such element which distinguished the French theocratic king from any other in medieval Europe: none had that distinguished ancestry and an oil of such hallowed origin as the rex christianissimus. It was not mere chris, but oil brought on a dove from heaven on the occasion of Chlovis’ baptism. Other kings had to go to the apothecary to buy their oil, the French king alone had his brought from heaven’ (Ullman, 1961, 201).

12 The language of its inhabitants was langue d’oc rather than langue d’oil and in many ways it was culturally and economically more connected to Catalonia and northern Italy than it was to the region north of the Loire (Given, 1990, 2008). Given notes ‘the French monarchy, however, had little to do with Languedoc since the break-up of the Carolingian empire. In the middle of the twelfth century the Capetians began to show a renewed interest in the affairs of the south. But when the wars to end Languedocian autonomy began in 1209, the French monarchy and the various southern lordships had behind them only a relatively brief history of important political contacts’ (Given, 1990, 39).

13 Goldsmith notes: ‘[t]he contrast between the wealth and institutional strength of the church in the north of France and the church in the Midi was long standing’ (Goldsmith, 2003, 238). Le Roy Ladurie found that the peasants of the region were often ignorant about many details of the Christian faith: ‘Sometimes people failed to fast through ignorance, I do not know what the fast days of the Church are, apart from Lent and Friday, said Guillaume Baille, a migrant shepherd of Montaillou’ (Ladurie, 1978, 314). The Church was distant to ordinary people and while major rites of passage like marriage were solemnized by a priest many rituals such as confirmation were neglected and ignored by ordinary people (Ladurie, 1978, 311–313).

22 According to this view, the Cathars held that the devil had created the material world and that Christ could not therefore have been born in human flesh and have suffered on the cross. Following the writings of contemporaneous Cistercian chroniclers, historians saw the Cathars as developing a church-like organization in opposition to the Catholic Church. See, for example, Oldenbourg (1961, 28–81), Lambert (1998), Hamilton (1999), and Barber (2000). This view sees the Cathars as decisively influenced by the Bogomil heresy which arose in Macedonia in the tenth century.

23 Traditionally the term ‘good men’ has been translated as perfecti as this was the term used to describe initiated or ‘perfected’ members of the Cathar Church. Pegg notes ‘What is so intriguing about the epithet “good man”, bons ome belonging to a person deserving of honor, and transcribed at Saint-Sernin as bonus homo or less frequently as prosbus homo, is that “good man” was the title adopted by any Larraugais man in situations circumscribed by courtesy’ (Pegg, 2001, 95).

24 For example, ‘Pons Estotz’, while confessing that he first believed in the faith of the bons omes in 1215, managed, somehow or other, not to listen to a word the good men said, or even to what anyone else might have said about them, until 1233, when, much to his surprise, he finally heard the good men say “that God did not make visible things, that the sacred Host is not the body of Christ, that baptism, like marriage is not salvation, and the bodies of the dead will not be resurrected”… Pons Estotz swore that “when I heard the heretics saying these errors, I left their faith at once” (Pegg, 2001, 75). Similarly Ladurie found that heretical beliefs could co-exist with orthodox ones and that individuals could move back and forth between heresy and orthodoxy (see Ladurie (1978, 311–313) and Ladurie (1978, 324–325).

Please cite this article in press as: Johnson, N.D., Koyama, M. Legal centralization and the birth of the secular state. Journal of Comparative Economics (2013), http://dx.doi.org/10.1016/j.jce.2013.08.001
There was no evidence of an organized Cathar church. Rather, and in a manner completely consistent with our model, the inquisitors created the image of the organized heresy they expected to encounter and in so doing they hardened the boundaries between orthodox Catholicism and heresy as ‘those who still believed in the heretics now did so with a new awareness about themselves: a self-consciousness that was, ironically, made precise and clear through the very men who wished to punish them, that is, through the two friar inquisitors’ (Pegg, 2001, 130). Discontent with the church hierarchy and church corruption was common throughout the middle ages and often led to dissenting belief (Leff, 1961, 1967). It was the growth of the medieval state, however, that criminalized this dissent and turned it into heresy.\(^\text{26}\)

Proposition 2 suggests that such a redefinition (or reduction in toleration—lower \(\tau\)) is more likely to occur if the region being absorbed into the centralizing state has a distribution of religious beliefs that lacks a clear sense of identity separate from that of the center. This describes the ‘good men’ of the Languedoc in the thirteenth century. To use the language of the model, the distribution of religious belief in the Languedoc was relatively non-cohesive (large value for \(\sigma\)) and non-deviant (low value for \(\mu_1 - \mu_2\)). It follows that the persecution of heretics in Languedoc can be seen as a consequence of the centralization of the French monarchy and the increased centralization of the Catholic Church. These two developments led to local religious practices being redefined as heretical. In the period before the crusades the good men of Languedoc seem to have seen themselves as orthodox Christians.\(^\text{27}\) Persecution turned them into Cathar heretics.

### 4.2. The reformation, legal centralization, and birth of religious toleration in France

The growth of state power enabled medieval monarchs to eradicate deviant religious practices as a way of enhancing their own legitimacy. Furthermore, consistent with our model’s predictions, because of the amorphous nature of religious belief in the south of France during the thirteenth century, the centralizing state faced little incentive to secularize legal institutions. Instead, it chose to restrict tolerance and, in the process created a ‘persecuting society’.\(^\text{28}\)

The Protestant Reformation disrupted this political-religious equilibrium and led to the spread of different religious doctrines. According to our model, an increase in the variance of religious beliefs within a single population initially leads to religious persecution, but also raises the costs of religious intolerance and reliance on religious legitimation. In the long run, therefore, this process creates incentives for the ruler to enforce rules granting greater religious toleration (Proposition 1). However, at the same time as the Reformation was leading to the diffusion of new beliefs concerning the relationship of individuals to God and the role of the Church, European monarchs were also occupied with imposing greater legal centralization over the disparate regions which they nominally ruled. We will show how this process of legal centralization initially intensified religious persecution until the costs of enforcing religious conformity became so great that they forced the French monarchy to grant a measure of religious toleration.

In 1500 there was little organized heresy in France (Cottret, 2003). From the 1520s onwards, the Reformation in Germany and Switzerland began to filter across the border. Renaissance rulers lacked knowledge or information on the religious beliefs of their populations. They therefore reacted to the Reformation by persecuting more individuals. The French monarchy executed the first heretic for ‘Lutheranism’ in 1524.\(^\text{29}\) The most important high court in the country, the Parlement of Paris, tried several other individuals for adhering to the Lutheran heresy in the 1520s but the overall numbers remained small (Salmon, 1975; Monter, 1999).

The French monarchy in the 1520s lacked the institutions necessary for investigating and prosecuting large numbers of religious dissenters. Religious courts responsible for enforcing conformity existed where local authorities were able, or motivated, to make them. In practice, beyond the seats of royal power, local officials typically took no action against suspected heretics.\(^\text{30}\) Papal inquisitors fervently pursued heretics in some regions, but were also often venal and ineffective. Monter observes that ‘[f]rom top to bottom from pope to bishop, church courts and canon law simply could not cope with the early Reformation’ (Monter, 1999, 45). Some parlements were known to be lenient to Protestantism, while others were known to take a hardline position. However, the parlements everywhere struggled to enforce royal authority especially ‘in the more distant parts of their jurisdictions’ (Greengrass, 1987, 38).

This began to change in the late 1530s. Francis I (1515–1547) pursued a policy of state-building and legal centralization. The edict of Villers-Cotterêts of 1539 imposed a set of common legal standards, including Roman canon law, across all regions of France (Mentzer, 1984). It was followed by the Edict of Fontainebleau in 1540 which specifically addressed the problem of heresy. The Edict of Fontainebleau:

\(^{26}\) Our framework is also applicable to other medieval heresies. Robert E. Lerner has demonstrated that the most important heresy of the fourteenth century, the Heresy of the Free Spirit was largely a construction of inquisition and that those convicted as belonging to this heresy had a range of beliefs that emphasized mysticism and apostolic poverty (Lerner, 1972). Forrest (2005) depicts in detail the construction of an apparatus for detecting heresy in England in the early fifteenth century. He observes that persecution created ‘a new community of the orthodoxy’ and ‘heresy and anti-heresy were produced co-dependently’ (Forrest, 2005, 235–236).

\(^{27}\) For example, the perfecta Rixende Buassan (later burnt), who was reported to have heard mass regularly, and to have made offering to the priest at the altar in the church of Sorèze (Arnold, 2001, 160).

\(^{28}\) For example, the perfecta Rixende Buassan (later burnt), who was reported to have heard mass regularly, and to have made offering to the priest at the altar in the church of Sorèze (Arnold, 2001, 160).

\(^{29}\) At this stage, though ‘the authorities had often little conception of what the term ‘Lutheran’ actually meant’ (Greengrass, 1987, 11).

\(^{30}\) Knecht describes the actions of the regional parlements as ‘uneven’, commenting that in ‘Dauphiné only eleven cases of heresy in seven years were reported and only light sentences given. Yet there were certainly Protestant groups in that province’ (Knecht, 1982, 401).

Please cite this article in press as: Johnson, N.D., Koyama, M. Legal centralization and the birth of the secular state. Journal of Comparative Economics (2013), http://dx.doi.org/10.1016/j.jce.2013.08.001
entrusted to all royal judges (i.e. members of the sovereign courts, baillis, sénéchaux, prévôts, etc.), ‘indifferently and concurrently’, the right of enquiry (inquisition) in respect of all persons, lay and ecclesiastical, except clergy in major orders. After prosecution, suspects were to be sent immediately to the sovereign courts for sentencing, regardless of any privilege or franchise they might claim. All feudal lords (hauts justiciers) were enjoined on pain of deprivation of their jurisdiction to carry out searching enquiries within their respective areas and to refer suspects to the king’s judges’ (Knecht, 1982, 398).

The Edict of Fontainebleau asserted secular control over heresy, making its suppression the responsibility of the regional parlements; these alone were permitted to use torture and to pronounce capital sentences (Mentzer, 1984, 45). It also represented a significant movement towards local standardization and it asserted royal control over the regional courts (which had to report to the king every six months) while at the same time granting those courts authority over religious affairs. Furthermore, for the first time, the Edict identified heresy with sedition and thus defined it as a crime against the state as opposed to the church (Sutherland (1980, 34) and Roelker (1996, 207)). Those church courts that continued to exist were increasingly supervised by royal officials (Mentzer, 1984, 10). At the same time, the definition of heresy was standardized in such a way that made it easier for the courts to convict suspects. It was shifted away from one based in theology (i.e. a judgement of the orthodoxy of one’s beliefs) towards an assessment of whether or not an individual’s visible behavior implied that they were heretical, as the latter was easier to ascertain. In 1544 the king published a list of prohibited books prepared by the Sorbonne and in 1545 ‘special commissions of parlementaries to seek out heresy in particular regions began to function’ (Roelker, 1996, 182).

These royal edicts created an official ‘engine of repression’ that could be applied across the country (Roelker, 1996, 211). Proposition 2 enables us to analyze the consequences of legal centralization. It predicts that when a common dogma is imposed on regions which have both cohesive and deviant distributions of beliefs large numbers of individuals will initially be categorized as heretics. However, the high costs of suppressing religious difference will eventually force the state to widen the bounds of religious toleration.

This is consistent with what took place in France in the first part of the sixteenth century (for a chronology see Table 1). Using trial data collected by Monter (1999) we illustrate the effect the creation of this engine of judicial repression had on the both the number of individuals executed for heresy and the geographic scope of these heresy trials. Figs. 9–11 show both the dramatic increase in heretic executions after 1540 and the accompanying increase in their geographic spread. Before the Edict of Fountainebleau, most executions occurred either in Paris or Provence. After 1540 almost all the regions in France began to execute heretics. These figures suggest strongly that the edicts of Francis I were not merely pieces of paper; they were actually obeyed.

The efforts of the King and the Parlement of Paris increased the geographic reach of the French centralized state as ‘inquisitors of the faith’ were appointed and sent out into the provinces in order to seek out and detect heretics (Roelker, 1996, 208). Fig. 12 depicts this increase at a local level by showing the geographic location of heretic executions in the généralité.
Fig. 10. Executions of heretics by généralité, 1540–1549.

Fig. 11. Executions of heretics by généralité, 1550–1560.

Fig. 12. Distance of heresy executions from high court in Provence located at Aix-en-Provence. Squares represent executions before 1540. Triangles represent executions after 1540. Multiple executions may be represented by one symbol.
of Provence. Squares show where executions occurred before 1540 (each symbol may represent multiple executions). Triangles show the location of executions after 1540. The Figure makes clear that before legal centralization in 1540, executions clustered closely to the primary seat of legal power for the crown in Aix-en-Provence. However, after 1540, executions occur much further away from the high court. In effect, the scope of royal authority, as measured by the willingness and ability of local courts to execute heretics, increased after 1540.

Fig. 13 shows similar data for all the généralités of France between 1523 and 1560. Each point in the scatter plot represents the distance of a trial from its regional high court (parlement). The plotted line is a kernel weighted local polynomial regression of distance of trial on year, or more simply, the average distance of trials from their high courts in that year. The data show that the centralizing reforms of Francis I between 1538 and 1540 were responsible for increasing both the number and geographic scope of heresy executions.

The increase in legal capacity associated with these edicts resulted in a tremendous growth in heresy trials, but it was not enough to suppress the Protestants: the number of heretics continued to grow throughout the 1550s. The campaign against heresy continued in the reign of Henri II (1547–1559) with the creation of the Chambre Ardente – a special court devoted to prosecuting cases of heresy – and the Edit of Chateaubriand in 1551. These marked the high point in the judicial repression of heresy.

This intensified persecution became increasingly ineffective, however; it served to strengthen the faith of Protestants and encouraged them to organize. Influenced by the rise of Calvinism in Switzerland, they began to develop a distinct corporate identity as Huguenots. As Kelley (1981, 92) observes, ‘... out of transient and sporadic dissidence the “so-called reformed

---

Table 1
Brief Chronology of the French Reformation, and the rise of toleration.

<table>
<thead>
<tr>
<th>Year</th>
<th>Event</th>
</tr>
</thead>
<tbody>
<tr>
<td>1523</td>
<td>First heretic executed for ‘Lutheranism’ in France</td>
</tr>
<tr>
<td>1534</td>
<td>Affair of the Placards marks intensification of repression</td>
</tr>
<tr>
<td>1540</td>
<td>Edict of Fontainebleau; imposes secular and royal control over heresy trials</td>
</tr>
<tr>
<td>1547</td>
<td>Creation of the Chambre Ardente, a specialist court to sentence heretics</td>
</tr>
<tr>
<td>1560</td>
<td>Conspiracy of Amboise fails</td>
</tr>
<tr>
<td>1562</td>
<td>Edict of January promises toleration to Protestant</td>
</tr>
<tr>
<td>1562</td>
<td>Beginning of the French Wars of Religion</td>
</tr>
<tr>
<td>1572</td>
<td>St. Bartholomew’s Day Massacre</td>
</tr>
<tr>
<td>1598</td>
<td>Edict of Nantes issued. Protection of civil liberties and toleration for Protestants</td>
</tr>
<tr>
<td>1685</td>
<td>Revocation of the Edict of Nantes</td>
</tr>
<tr>
<td>1724 onwards</td>
<td>Relaxation of the enforcement of the Edict of Nantes</td>
</tr>
<tr>
<td>1787</td>
<td>Edict of Versailles grants toleration to non-Catholics</td>
</tr>
</tbody>
</table>

---

34 The Edict of Châteaubriant of 27 June 1551 was ‘un vrai code de persécution’ (Sutherland, 1980, 44). Many members of parlement themselves became suspects under its auspices. It included twenty articles ‘covering in exhaustive detail all matters of censorship, the possession, production, sale, and dissemination of religious works; the rigid control of printing, and the inspection of bookshops. But even more vexatious and dangerous were the articles (27–33) relating to informing, which became mandatory (Sutherland, 1980, 46).
religion” had created a coherent movement and an almost complete corporate organization. The Huguenots, unlike the heretics of the thirteenth century, could call upon an international Protestant movement for intellectual and spiritual support, and perhaps more importantly still, they could employ the printing press to celebrate those who died as martyrs (Kelley, 1972; Parker, 1993).

Proposition 2 suggests that attempts to impose legal standardization in the face of large religious differences (high deviancy and high cohesiveness) are likely to be counterproductive. Monter (1999, 120) observes, the ‘... great heresy hunt begun in the mid-1540s had strained the resources of the Paris Parlement to their limits. Accused heretics crowded their jail and their workloads. Moreover, some of these cases presented serious professional difficulties. Heretics were not easily or comfortably judged by men who had mastered every nuance of Roman Law but were untrained in theology’. Despite the imposition of royal control over the regional and church courts, enforcement of the laws against heresy continued to vary from region to region. In the south particularly, landowners sympathetic to Protestantism sheltered and failed to try individuals suspected of heresy.

The failure of the policy of persecution to destroy Protestantism was apparent to French rulers by 1560. The death of Henri II in 1559 prompted a political crisis. A section of the nobility led by Louis de Bourbon, Prince of Condé, had become Protestant and opposed the policy of persecution. Henri’s widow, the queen regent, Catherine de Medici recognized the costs of maintaining the campaign against heresy and decriminalized Protestant beliefs. She issued the Edict of Amboise in March 1560, which promised a general pardon for all heretics so long as they returned to the Catholic fold. It aimed at avoiding a massive effusion of blood and was an acknowledgment that the eliminationist policy initiated by the Edicts of Paris and Fontainebleau was self-destructive (Sutherland, 1980, 105). However, it did not offer the Protestants freedom of worship, which is what their leaders now demanded. This came with the Edict of January issued at St. Germain in 1562 which was the first Edict to offer some form of religious toleration. This recognized, for the first time in French history, the right of private individuals to practice their own faith in their own homes: it was the ‘first indication of the possibility that citizenship and the political allegiance of subjects to their king and country might be consistent with a diversity of religions in the state’ (Zagorin, 2003, 91).

This account of the evolution of legal tolerance for religious belief between 1523 and 1562 in France is consistent with our model. The Reformation increased the variance of religious beliefs (σ). Rulers who attempted to maintain religious orthodoxy (hold τ constant) consequently found themselves persecuting large numbers of individuals. In France the attempt to impose legal standardization over a large and diverse kingdom exacerbated and amplified the effects of the Reformation. It meant that the French campaign against Protestantism was particularly costly and bloody compared to other European countries. The costs of religious persecution, furthermore, eventually forced the French monarchy to abandon intolerant policies. In the case of the Reformation in France, legal centralization was accompanied by greater toleration because unlike the ‘good men’ of the thirteenth century who lacked a strong sense of identity, the Protestant movement was well defined and distinct from the Catholic Church. To use the terms of our model, the sixteenth century Protestant movement was more cohesive (low σj) and deviant (high (μj – μm)2) than the thirteenth century Cathars. Thus, the mixture of the distributions of religious belief faced by the centralizing state resulted in a much higher variance in belief than for any individual region.

The Edict of January was important as the first acknowledgment of religious toleration in French history. It accompanied similar edicts in the rest of north-western Europe such as the Peace of Augsburg in Germany in 1555 and the Act of Uniformity in England in 1558. The Peace of Augsburg recognized the right of Lutherans to practice their religion in those German states ruled by Protestants and the Act of Uniformity recognized the right of Catholics in England to private worship. However, in France at least, there was a sense in which official secularization came too late as civil war between Protestant and Catholic factions soon followed. The war was accompanied by large-scale killings of Protestants by Catholics, culminating in the St. Bartholomew’s Day massacre in 1572 which resulted in between 5000 and 20,000 deaths according to modern estimates.

The civil war ultimately resulted in a victory for the Protestant leader Henri of Navarre, but in order to be accepted as king of France as Henri IV he had to credibly commit his side to maintaining a secular legal stance by converting to Catholicism. In order to conciliate the Protestants, Henri IV (1589–1610) issued the Edict of Nantes in 1598 which provided religious toleration for the Huguenots and ensured their position within France for roughly the next 100 years. The Edict of Nantes solidified the new equilibrium that the legally unified kingdom of France started moving towards after the reforms of Francis I in the 1530s.

4.3. Our model applied to other states

The experience of other European states are largely consistent with our model. In England, the historical rise of religious toleration was a gradual process, or in the words of MacCulloch, ‘crabwise’ (MacCulloch, 2003, 679). This is consistent with our model. As in France, the 1530s saw a process of political centralization which historians have termed the ‘Tudor
Revolution in Government’ (Elton, 1953). At the same time, a growing number of individuals were executed for heresy in the 1530s and 1540s. This culminated in the reign of Mary I (1553–1558), who had more than 280 Protestants executed in less than five years. However, the costs of these trials put severe strain on Mary’s government (in terms of our model $\tau < \tau^*$), and, on the accession of Elizabeth I (1558–1603), a decision was made to end heresy trials because it was too costly to criminalize the beliefs of such a large proportion of the population (McGrath, 1967, 1990). The Act of Uniformity of 1560 made Church attendance on Sunday compulsory and non-attendance punishable by a sizable fine. It did not enquire deeply into an individual’s actual beliefs and so-called Church Papists – those who conformed to Anglican worship but considered themselves Catholic – were tolerated (Walsham, 1993). Like the Edict of Nantes, this represented a widening of the bounds of toleration even if it heavily restricted freedom of worship. The last individual to be burnt for heresy in England died in 1612. Catholic priests and those who aided them continued to be executed for treason until 1681 and a severe campaign of repression against Protestant dissenters was launched in 1662. The Act of Toleration of 1689 represented a further extension of the bounds of toleration – it granted toleration to all Protestants but maintained civil disabilities on Catholics until 1828.37

The Netherlands also experienced a severe religious persecution in the mid-sixteenth century associated with attempts by the Habsburgs to impose more centralized and direct control on the province which eventually provoked a rebellion colored by religious tension. The success of the Dutch revolt against Habsburg rule brought to power a Calvinist party that believed in religious persecution. However, because of the Reformation and civil war, there was now a large number of people in the new Dutch Republic who did not subscribe to Calvinism (MacCulloch, 2003, 367–378). As a result, the Calvinist party had to accept that it was infeasible to persecute Catholics or other Protestants and, out of the resulting compromise, a form of de facto religious toleration gradually emerged during the seventeenth century (Zagorin, 2003).

In Germany the Wars of Religion of the mid-sixteenth century were brought to an end by the Peace of Augsburg (1555) and the maxim cuius regio, eius religio. This recognized Lutheranism and Catholicism (though not Calvinism) as legitimate religions and left it to each local ruler to decide which would be the official belief of his region. This was quite different to the situation in France. The Peace of Augsburg made official the policy that each local jurisdiction could set its own, relatively narrow, standard of orthodox belief. As one historian puts it ‘the license was for the princes rather than the subjects’ (Sutherland, 1984, 159). From the point of view of our theoretical framework, this devolution of legal authority is highly consistent with what one would expect in the presence of a weak central state. This is precisely how historians describe the situation. The Holy Roman Emperor was too weak to impose religious conformity at an empire-wide level, but this was a reflection of the fact that real sovereignty in Germany was increasingly held by the statelets and principalties within the empire. As Sutherland observes: ‘toleration could not simply be proclaimed. No government could – successfully – be more tolerant than the governed and tolerance was emphatically not the spirit of the age. Toleration required . . . an exceptionally strong secular power, such as Henri IV of France’ (Sutherland, 1984, 155).

Like the German states, Poland provides an interesting counterexample to France, but one that also supports the predictions of our model. Sixteenth century Poland has been heralded as ‘a state without stakes’ – an almost unique example of religious toleration and pluralism in a period of ubiquitous religious conflict (Tazbir, 1973). Catholic Poland did not persecute Protestants, nor did it even prosecute Anti-Trinitarians who were denied toleration almost everywhere else. And, although it was not true that Poland was a state entirely without stakes – a woman was burnt to death in 1539 for denying the divinity of Christ and attempting to convert to Judaism—it certainly did not go through a campaign of religious persecution on the same scale as other regions in Europe.38

What makes the example of Poland consistent with our framework is that the Polish state did not go through the same process of state centralization which accompanied religious prosecution in England, France, and the Netherlands. The Polish monarchy under Sigismund I (1506–1568) passed edicts against heresy, but, unlike the actions of the kings of France, these were ineffective because the Polish state was decentralized and the monarchy – circumscribed by the liberum veto which enabled the nobility to veto royal policy – comparatively weak (see Roháč, 2008).39 Religious policy remained the responsibility of ecclesiastical courts with limited jurisdiction – there was no Polish equivalent to the Edicts of Paris and Fontainebleau. The power of the Polish aristocracy was such that no nobleman was sentenced to death for his beliefs throughout the period and ‘[n]o parliament could agree to any law invoking the death penalty for noblemen. This prevent the state from even being able to exercise the bounds of toleration – it granted toleration to all Protestants but maintained civil disabilities on Catholics until 1828.37

37 Murphy (2001) notes that it was less liberal than James II’s Acts of Indulgence (1687 and 1688) or the policy of Oliver Cromwell during the Protectorate. MacCulloch further notes, even ‘when John Locke published his celebrated Letters concerning toleration in the aftermath of England’s Glorious Revolution, he still excluded Catholics and atheists from his proposals’ (MacCulloch, 2003, 670). The last person to die for their religious beliefs in the British Isles was a student in Edinburgh for blasphemy in 1697 (Coffey, 2000, 200). Coffey notes: ‘[i]f 1689 marked the end of the persecutory society, it did not mark the beginning of the secular state’ and throughout the eighteenth century anti-Catholic legislation continued to be passed and enforced (Coffey, 2000, 201).

38 See Teter (2006, 44) for details of the execution of Katarzyna Malcherowa in 1539.

39 Furthermore, Poland unlike other western European countries was never religiously homogeneous, and consequently, Catholicism played a much less important role in legitimizing the state in Poland than elsewhere in Europe. The state of Poland–Lithuania extended over a vast territory, and although its rulers were Catholic, its population included significant Greek Orthodox, Armenian, Jewish and Muslim minorities (see Tazbir, 1973, 31). The religious diversity of medieval Poland meant that even prior to the Reformation the hardline Church position against heresy was not enforced. Consequently, the clergy in Poland advocated toleration and peaceful coexistence; in terms of our theory: the Polish-Lithuanian commonwealth survived in part because the bounds of toleration were set fairly widely

40 Also see Lecler (1960). Transylvania is also a case in point.
expelled on pain of death in 1658. While some of these left the country or converted, however, many ‘… continued to live in Poland without abandoning their faith … One of the reasons was the general slackness of the executive, which never enforced the anti-Arian laws very strictly, especially against members of the nobility who always enjoyed certain immunities’ (Tazbir, 1973, 198). As Davis notes: ‘The nobility believed what they wished, and protected whom they liked. The bourgeoisie and the Jews were secure within the framework of their autonomous estates … the ‘Golden Freedom’ of the nobility, proved an obstruction to efficient government and to religious fanaticism alike’ (Davies, 1981, 200). As in Germany, this situation was not one of modern religious toleration but a sorting equilibrium in which local nobles were able to practice the religion of their choice because of the weakness of the state.

Finally, our interpretation is supported by the fact that other large empires with strong states and heterogenous populations have also been characterized by surprisingly high levels of religious toleration. Notable among these are the Persian empire, the Roman empire and the empire of Genghis Khan and his successors. Our model is perhaps best illustrated, however, by the Islamic empires of the Middle East which were relatively tolerant for most of their histories.

Islam promised toleration for people of the Abrahamic religions, and historians agree that their attitude of comparative religious tolerance (and indifference to say, the disputes between Monophysites and Dyophysites) enabled Arab armies to make rapid conquests throughout the Eastern Mediterranean. Our theory suggests that Islamic societies were tolerant because the Caliphate was large and encompassed many groups with heterogenous beliefs. The Islamic Caliphate had to be religiously pluralistic for pragmatic reasons. Even within the Arabian peninsula, Muhammad and the early caliphs received support from Jews and later Christians and Zoroastrians who ‘stood firm over their own faiths but were friendly disposed towards the Prophet and his new creed’ (Bosworth, 1982, 41–42). What was later known as the pact of Umar a ‘contract in which the non-Muslims agree to a host of discriminatory regulations in return for protection’ was in fact a rationalization of these ad hoc arrangements (Bosworth, 1982; Cohen, 1994).

The explanation for state secularization suggested by our theory is supported by the fact that as the Arab armies continued to conquer new regions their political leaders found themselves forced to extend toleration first to ‘people of the Book’, then to Zoroastrians, and then, during the conquest of Sindh in northern India in the eighth century, to polytheistic Hindus.

This tradition of religious pluralism was continued by the Ottomans when they conquered Eastern Europe where it became known as the millet system (Runciman, 1970; Bruade, 1982). Non-Muslims had dhimma status: they could not proselytize and had to pay additional taxes, and sometimes had to wear special, distinctive, clothing, but they were largely free from persecution and they were allowed to continue to use their own legal systems (Bruade and Lewis, 1982). The Ottoman empire was much more tolerant of religious diversity than almost any European state through the early modern period, but it did not move towards the liberal model of religious freedom that eventually emerged in Europe.

5. Discussion and concluding comments

What of the many alternative explanations for the rise of the secular state that do not portray legal centralization as playing a central role? Traditionally, religious toleration has been seen as a product of the Enlightenment and is associated with the ideas and writings of the French philosophes (Kamen, 1967, 2000). However, the historical literature we have discussed in this paper demonstrates that the origins of religious toleration lie in the Reformation in the sixteenth century, a time when almost no one was arguing in favor of religious liberty.

The Edict of Nantes along with the experience of religious toleration in the Dutch Republic helped to inspire many of the most powerful intellectual voices for greater religious toleration such as John Milton (d. 1674), Roger Williams (d. 1683), Pierre Bayle (d. 1706), and John Locke (d. 1704). While almost no sixteenth century thinker openly defended the idea of complete religious freedom, in the seventeenth century increasingly radical arguments in favor of greater tolerance were developed. Locke argued that faith could not be compelled and argued for complete religious toleration for Protestants, but

---

41 The Persian empire during the Achaemenid period (559–330 B.C.) similarly practiced religious tolerance (Masrooi, 1999, see). It appears to have done so in part to order rule a large empire with a number of diverse religions. Interestingly, later Persian empires such as the Sassanid empire (224–651 A.D.) which ruled much smaller, and more cultural homogenous territories, practiced religious persecutions.

42 Furthermore, once the Caliphate had been established the economic role played by minorities in the Islamic Middle East and the special taxes that they paid helped to ensure that they would be tolerated, although, minorities were occasionally persecuted at various points in time (Cohen, 1994). Bernard Lewis notes: ‘[v]iolent persecution, forced conversion and banishment were rare, though not unknown. They usually occurred at times of stress and danger, when the Islamic world was threatened from within or without, by pestilence or famine, religious division or foreign invasion’ (Lewis, 1986, 124).

43 The granting of dhimmi status to Zoroastrians was justified because it was claimed that they had been granted a revelation from God but the Holy Books has not survived. However, ‘[w]hen the Muslims first acquired a foothold in the Indian subcontinent through the conquest of Sindh in the early eighth century, a situation arose similar to that of Persia; the teeming populations there could not be slaughtered en bloc, but how in the absence of any Quranic nass, could the pagan Hindus be assimilated to dhimmi status? On the evidence of Baladhuri’s account of the conquest of Sind, there were certainly massacres in the towns of Sindh when the Arabs first arrived’. But, eventually ‘[p]lace treaties were made with the local communities … on what had been the standard conditions during the overrunning of the Fertile Crescent and Persia’ (Bosworth, 1982, 43).

44 And, as Kuran (2004, 2010) has argued, this choice of legal system gave Christian and Jewish minorities and economic advantage once trade with western Europe opened up in the eighteenth and nineteenth centuries.

45 The failure of the Ottoman Empire to continue on the same path of secularization as its European neighbors can be at least partly explained by the gradual weakening of the Ottoman state that began in the seventeenth century (Balla and Johnson, 2000; Pamuk, 2004).

46 Castello wrote under a pseudonym. Jean Bodin’s argument in favor of religious toleration for other faiths (not for atheists) in the Colloquium Heptaplomeres was not published (see Kuntz, 1998).
nonetheless maintained that a state could persecute Catholics, Muslims, and atheists on secular grounds. Bayle went much further in arguing that, contra-Locke, a society of atheists could in fact function perfectly adequately. Building on this, Adam Smith would argue that ‘[s]ociety may subsist among different men, as among different merchants, from a sense of its utility’ (Smith, 1759, 100). This was the argument that allowed the American Founders to argue for complete religious freedom in establishing the First Amendment in 1791.

Our argument in this paper takes nothing away from the ideas of these advocates for religious toleration. However, important as these intellectual justifications for religious freedom were, they followed de jure increases in religious toleration that were driven not by Enlightenment philosophy, but by the centralizing legal reforms which accompanied the rise of the early modern state.

A second, influential explanation for the rise of religious toleration is the doux commerce thesis associated with Montesquieu, Hume, and Smith (see Hirschman, 1977). This argument suggests the growth of markets and trade increases tolerance by bringing together individuals with different customs and religious practices. This is undoubtedly true and has found support in recent empirical studies by economists. Nevertheless, the doux commerce thesis explains the gradual amelioration of religious hatreds and tensions, while religious toleration was a political achievement – one that took place in Europe at a time when religious hatreds remained strong. The evidence we provide suggests that any effect of doux commerce occurred after a political framework for religious toleration had already been established. The doux commerce hypothesis thus complements, and is not a substitute for, the thesis we have advanced in this paper.

The framework developed in this paper can provide one explanation for the inverse-u relationship between the centralization of state institutions and religious persecutions in medieval and early modern Europe. This argument has important implications since it implies, not only that legally centralized states played an important role in the secularization of policy, but also that contemporary states that are investing in legal capacity may initially experience increased levels of religious violence before the official bounds of tolerance are eventually relaxed. Zagorin (2003) makes the intriguing observation that no society was as intolerant of religious differences as was western Europe in the medieval and early-modern period, but paradoxically, it was also in western Europe that the idea of religious freedom was fully developed in the sixteenth century. As he writes, it ‘is only in Western society, nevertheless, and only since the sixteenth century because of the conflicts and debates between contending Christian churches, sects, and confessions, that there has appeared a massive body of writings by different authors exploring the problem of religious toleration from many angles and presenting an array of arguments on behalf of the principles of liberty of conscience, mutual toleration, and religious coexistence and diversity’ (Zagorin, 2003, xiii). According to our thesis, the great outburst of writings making the intellectual case for toleration, described by Zagorin, was preceded by the legislative work of European rulers, who were motivated less by lofty ideals than by their own enlightened self-interest. The secular state was born of necessity and compromise, a reluctant concession on the part of rulers of increasingly centralized states.

Acknowledgments

We are grateful to Metin Cosgel, Alexandra Mislin, Jared Rubin, Tyler Cowen, Joshua Trojniak, Timothy Shah, Mario Ferrero, Jean-Paul Carvalho, Bryony Reich and audiences at ASREC and the Washington Area Economic History Workshop for giving us feedback on earlier versions of this paper.

References


Alesina, Alberto, Reich, Bryony, 2013. Nation-Building. memo.


Clingingsmith et al. (2009) find that attendance of the Hajj increased tolerance and inter-faith harmony. They suggest that this is due to the contact made with other pilgrims from other parts of the world during the Hajj.

A related, argument suggests that religious tensions and conflicts diminish in intensity along with the overall trend towards secularization in society (Chadwick, 1975). This observation is highly plausible, but, contrary to the claims of Jordan (1932), does not apply to the early modern period when religious was, if anything, increasing.

49 This complementary relationship would only be strengthened, furthermore, by any beneficial effects on commerce due to the centralization of legal institutions by the state.


