EXPLORING STATE AUTHORIZATION STRINGENCY: A MACHINE LEARNING APPROACH

Jacob Fowles
Associate Professor, University of Kansas

ACKNOWLEDGMENTS

This research was supported by a research grant from the State Higher Education Executive Officers Association and Arnold Ventures. I am grateful for Erik Ness, Sean Baser, and the entire research team that endeavored to complete the initial data collection upon which this research is built. All errors are my own.

This paper is one in a series of reports coordinated by the State Higher Education Executive Officers Association (SHEEO) and supported by Arnold Ventures. Given increased public concerns about educational quality, the series is designed to generate innovative empirical research regarding state authorization processes and policies that can serve as a foundation for future research and policy in this understudied area. The views expressed in this paper – and all papers in this series – are those of its author(s) and do not necessarily reflect the views of SHEEO or Arnold Ventures.

INTRODUCTION

Today's economy makes clear that higher education is no longer relegated to being a luxury for the most privileged members of the upper class but a necessity for access to economic mobility and opportunity. Despite the continued increase in the costs associated with the pursuit of higher education—increases which significantly outpace inflation and growth in median wages—numerous studies demonstrate that the return on investment (ROI) of college attendance continues to make it one of the best investments an individual can make in their future.

However, the ROI is not guaranteed: Numerous lawsuits attest to the fact that some students enroll in and make significant tuition payments to institutions of higher education that fail to follow through on the promise to provide knowledge and skills deemed valuable by the job market. Further, the diversity of the higher education landscape makes it difficult, if not impossible, for employers to possess sufficient information about each institution to make informed judgments about the value of various degrees held by potential employees. These issues represent two important forms of information asymmetry in the market for higher education that have historically been utilized to justify governmental intervention—both to protect students from bad actors and to provide assurances to employers that degrees held by applicants provide meaningful and useful signals of applicant quality and ability.

Accordingly, governments at all levels have long been involved in the regulation of higher education markets, although numerous scholars point out that the state remains the primary actor in the higher education regulatory space. Scholarship across such academic domains as higher education, political science, public policy, and economics has devoted no small amount of attention to identifying, exploring, categorizing, and, to the extent possible given the difficulty in obtaining relevant data, evaluating these interventions. A dominant vein of that literature considers the role of state higher education governance, defined as “the different ways in which states have organized their respective systems of higher education, and authority patterns within those systems” (McLendon, 2003, p. 58). The so-called “modern” era of those systems was born in the aftermath of World War II, a time in which higher education underwent “massification,” shifting from a good consumed primarily by elites to one made available to an increasingly growing middle class, a movement that was catalyzed by the passage of the Servicemen’s Readjustment Act of 1944 (popularly known as the G.I. Bill), which included generous funding for returning veterans to pursue higher education (Thelin, 2011).

The central analytical paradigm in the scholarly literature on higher education governance was defined by the pioneering work of McGuinness (1994), which categorized states according to two dimensions: campus governance and statewide coordination. Campus governance refers to the way in which a state provides for the day-to-day management of its higher education institutions, while statewide coordination references the mechanisms utilized by the state to facilitate coordination and cooperation across those institutions (McLendon, 2003).

McGuiness distilled from these two dimensions a typology that recognizes three unique governance arrangements: advisory boards (possessing little regulatory power in either of the two dimensions), coordinating boards (possessing regulatory power with respect to statewide coordination but little in terms of campus governance), and consolidated governing boards (possessing regulatory power in both domains). This typology has gained significant traction in the literature, with a robust literature exploring both the antecedents as well as the impacts of the “holy trinity” of higher education governance types (McLendon 2003, Volkwein & Tandberg, 2008; Tandberg, 2013).

However, recent scholarship has suggested that this two-dimensional picture of state higher education management is incomplete in that it ignores the role of states as key actors as authorizers of institutions of higher education. Here, authorization refers to the necessity of a higher education institution to obtain an initial approval to commence operation within a state (“initial authorization”) and to continue demonstrating compliance with the individual statutes, regulations, and rules of the state with respect to any ongoing educational functions (“continued authorization” or “reauthorization”). Critically, the rise of for-profit and exclusively online higher education has afforded the topic of authorization some degree of increased visibility; however, it would be inappropriate to think of authorization as an issue relevant only within these contexts. Broadly speaking, the lack of understanding of the role of authorization within the broader higher education governance landscape has hindered our understanding of the nature of accountability between institutions and the states in which they function, especially given the recognition that authorization is best understood as a cross-cutting activity that may or may not respect the institutional boundaries identified in the literature (Tandberg, Brueckner, & Weeden, 2019). Generally speaking, states require institutions to obtain approval for engaging in regulated activities within the state. However, there exists a surprising degree of heterogeneity in terms of the scope of activities subject to regulation, leading Kelly, James, and Columbus (2015) to refer to the landscape of regulation of American higher education as

50 separate state regulatory regimes rather than one coherent and cohesive system. Common activities subject to authorization include online learning, experiential learning, marketing, and advertising within a state, as well as conferring degrees and non-degree credentials.

Further, states vary in terms of delegation of this authority. In some places, it is concentrated within a single bureaucratic agency, and in others, a variety of different agencies may share responsibility. As Tandberg et al. (2019) argue, the lack of systematic understanding about how these regimes operate is an oversight with potentially important ramifications for state efforts to increase and manage quality, access, and affordability within higher education.

Specification of Research Questions

To that end, this research draws upon recent work to codify and categorize the stringency of state higher education authorization regimes (Ness, Baser, & Dean, 2021) and, building on those foundations, provides some preliminary evidence about the nature of state approaches to higher education authorization practice, policy, and procedure, by addressing two basic, exploratory research questions:

1. **Can unsupervised machine learning techniques identify distinct clustering of states with similarities in approach to higher education authorization? If so, what characterizes those similarities?**
2. **Are the economic, social, political, and institutional characteristics identified in the state higher education policy literature useful for understanding and explaining the assignment of states to authorization clusters?**

This paper proceeds as follows. For each research question, an overview and theoretical framework are provided. This is followed by a discussion of the empirical model, method, and datasets employed in addressing each question. Next, the empirical results are presented with some discussion about the relative fit of the models to the data and substantive interpretation of the findings. Finally, the paper provides some overall discussion and conclusions for both research questions, concluding with some suggestions for particularly fruitful avenues for future research.

RESEARCH QUESTION 1

Introduction and Theoretical Framework

Regulatory stringency is a topic that has received some empirical attention across policy domains and academic disciplines. One key hindrance to such efforts is the lack of a consistent approach to conceptualizing and measuring stringency itself, leading Judge-Lord, McDermott, and Cashore (2020) to characterize the literature on regulatory stringency as "vibrant but confusing." In that work, the authors review the literature and identify at least five distinct approaches to characterizing stringency, with vertical notions of variation (ranking regulatory environments from most to least stringent based on evaluations of the intensity of regulation) and horizontal notations of variation (categorizing stringency according to the breadth or width of the scope of regulation) representing perhaps the two most dominant paradigms. Empirically, both approaches are often operationalized through the process of capturing qualitative variables representing the presence or absence of particular provisions in regulatory law or administrative practice and then producing an additive index that combines the individual indicators into a single composite measure. For example, Bott and Kozluk (2014) develop a country-specific measure of environmental policy stringency by collecting and coding the qualitative information contained in laws and regulations governing the energy sector, collecting such information as prices of CO2 allowances, tax rates associated with various pollutant byproducts of energy production, and governmental expenditures for research and development into renewable energy technologies, among others. The authors engage in comparative analyses of these features across countries, scoring stringency based on observed relative differences. These individual metrics are then aggregated to produce a country-specific score.

In other words, a key conceptual issue in understanding regulatory stringency is the problem of multidimensionality: how to reduce the complex and multifaceted nature of regulatory policy to a single generalizable, valid, and reliable metric. Brunel and Levinson (2013) note that to solve the multidimensionality issue associated with measuring regulatory stringency, researchers have historically adopted one of two approaches: either narrowing the breadth of stringency by focusing on a few stringency measures chosen for ease of comparison or by utilizing broad, comprehensive indices. The problem inherent to the former approach is generalizability: How can one be sure that the narrow conceptualization of stringency is meaningful beyond its immediate context? The problem with the latter approach is interpretation and comprehension: How can a measure that reflects such broad, varying contexts and idiosyncratic nuances be of pragmatic use and intellectual value? This tension underpins the search for identification and evaluation of theoretical frameworks explaining such complex dynamics.

However, recent developments in the field of unsupervised learning offer the potential for overcoming these limitations. Drawing upon that work, I apply Gaussian Mixture Modeling (GMM) to the dimensions of regulatory stringency developed by Ness, Baser, and Dean (2021) to assess the extent to which individual state authorization regimes can be grouped into distinct clusters based on stringency across substantive authorization domains. In their foundational work, Ness, Baser, and Dean (2021) measure and categorize the stringency of four distinct dimensions of a state's posture regarding higher education authorization: organizational and governance, academics, consumer protections, and student outcomes. They accomplish this by gathering data on 41 individual metrics for the agencies tasked with state authorization across the 50 states. The individual metrics are all coded on a 3-point scale, with 0 corresponding to the absence of utilization of a particular metric, 1 corresponding to utilization of a metric, and 2 corresponding to not only mandatory utilization and reporting requirements for a metric but also the utilization of a threshold of acceptability for the metric or additional stipulation (Ness, Baser, & Dean, 2021). Given that states often employ multiple agencies and processes in the context of authorization, state-level measures were constructed for each domain using additive, process-level averages within states.

Table 1 presents definitions and descriptive statistics for the individually identified dimensions of stringency. The first column conveys the description of the dimension. The second column contains the number of unique metrics (indicators) collected for each dimension (each of which is scored on a scale of 0-2). Columns 3-5 contain descriptive statistics for each dimension, containing the median, minimum value, maximum value, and interquartile range observed in the dataset for the 50 states and Washington, D.C.

Table 1: Descriptive Statistics of Stringency Dimensions

Dimension	Unique Metrics	Median	Minimum	Maximum	IQR
Organizational & Governance	11	13.0	1.0	22.0	5.4
Academic	13	17.5	0.0	24.5	7.5
Consumer Protection	10	11.0	0.0	17.0	4.4
Student Outcomes	7	1.0	0.0	10.0	3.8

As *Table 1* conveys, there exists a significant degree of heterogeneity across the individual states in terms of stringency across the identified substantive dimensions. Interestingly, the interquartile range, which conveys the spread of the middle 50% of values, is generally consistent across dimensions with one exception: academic metrics.

Ness, Baser, and Dean (2021) utilize various metrics to generate additive stringency metrics, ranking state authorization agencies by the relative level of stringency in approach. This research builds upon but departs from that work in that the focus here is on searching for commonalities in the latent configuration of authorization stringency across states, rather than comparing the relative levels of stringency across states. Accordingly, one potential outcome of this research is to confirm the additive categorization approach taken by Ness, Baser, and Dean—implying that groups of states broadly align with respect to the relative levels of stringency with which various facets of authorization are managed—but other conclusions are possible. In other words, it is not a priori obvious whether the individual states consider stringency across the identified dimensions as complements, substitutes, or as wholly unrelated to each other. A cursory review of the literature suggests numerous competing theoretical and conceptual hypotheses that may be explored. The formative work of McCubbins and Schwartz (1984) argues that policymakers possess rational preferences for forms of oversight as a mechanism of regulation and performance management. Application of this concept to the current context lends credence to the idea that the substantive dimensions that Ness, Baser, and Dean (2021) identify could be viewed as substitutes: A high level of stringency in one dimension may provide spillover protections that render a high level of stringency in another unnecessary or inefficient.

Conversely, one can imagine states simply taking a one-size-fits-all approach to authorization stringency, implying a standardization of intensity of stringency across dimensions. Application of the rich policy diffusion and innovation literature from political science and public policy yields additional relevant context. For example, application of this literature implies the possibility of emulation and policy learning across states: If one state discovers a particularly efficient or effective configuration for managing authorization, other states may adopt a similar configuration, leading to clustering and alignment over time (Shipan & Volden, 2008). Adding additional nuance to this discussion is the emerging recognition that states may look only to others with meaningful similarities in underlying characteristics (political, economic, sociological, etc.) when considering innovations, thereby incorporating the influence of internal determinants into dynamic diffusion models (Grossback, Nicholson-Crotty & Peterson, 2004). Or perhaps multiple or indeed none of the above hypotheses hold, indicating the potential for a “garbage can” approach to authorization policies and practices where solutions are formulated on a more ad hoc basis and only loosely coupled with context (Cohen, March, & Olsen, 1972).

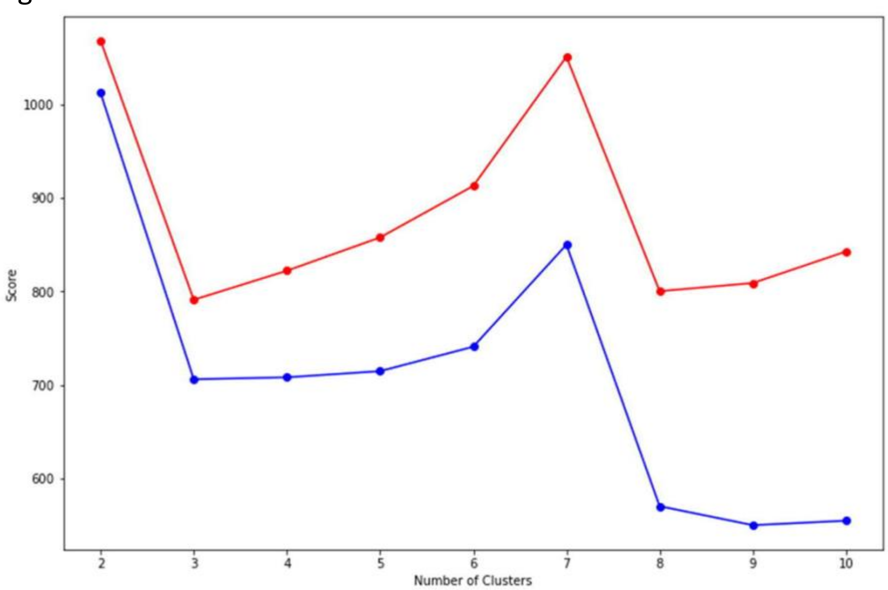
Methods and Model Specification

GMMs, as one technique in the family of unsupervised machine learning approaches, are particularly well suited to generating evidence to begin evaluating these competing hypotheses. Unsupervised learning approaches are particularly useful for exploratory data analysis because they are designed to discover hidden patterns or groupings in data with minimal input from researchers. GMMs, as a particular type of unsupervised machine learning model, assume a number of underlying and latent unique clusters, each of which is defined by a unique and latent underlying Gaussian distribution. The task at hand is, therefore, to utilize the k-dimensional data (k representing the number of unique characteristics available for each observation in the data) to estimate the optimal values for the latent parameters of this distribution by ensuring that each Gaussian fits the data points assigned to it, a task that relies upon the Expectation-Maximization algorithm (McLachlan & Basford 1988). A desirable feature of GMMs is that they are more flexible than more common clustering approaches such as K-means, in that they make comparatively less restrictive assumptions about the nature of the underlying clustering. As noted in the literature, one key assumption made by the researcher in employing GMMs is the choice of a number of unique clusters to be identified by the algorithm. Various proposals appear in the literature to guide that choice. Here, for simplicity, I simply rely upon an iterative trial and error process. That is, I try a range of different numbers of clusters and utilize Akaike information criterion (AIC) and Bayesian information criterion (BIC) to assess the relative fit of each estimated model (Burnham & Anderson, 2004).

In so doing, I initially estimate nine separate GMMs: The first one specifies two unique clusters, and the last specifies 10, with each model in between adding an additional cluster to the estimation procedure. For each iteration of this exercise, I calculate and record the AIC and BIC scores of the model. *Figure 1* below provides a graphical representation of the results of this exercise. The x-axis of the figure captures the specified number of clusters utilized in each estimated GMM, and the y-axis captures the AIC and BIC scores associated with that model.

Lower AIC and BIC scores indicate better-fitting models. Visually, moving from left to right, a steep (flat) downward slope between two points represents a significant (insignificant) improvement in model fit associated with the addition of an additional cluster. Upward sloping lines between two points demark a deterioration of fit. General guidance from the literature is to choose the number of clusters by identifying points of inflection where the slope of the line connecting the models experiences a significant shift from steep to flat.

Figure 1: AIC and BIC of Estimated GMMs*



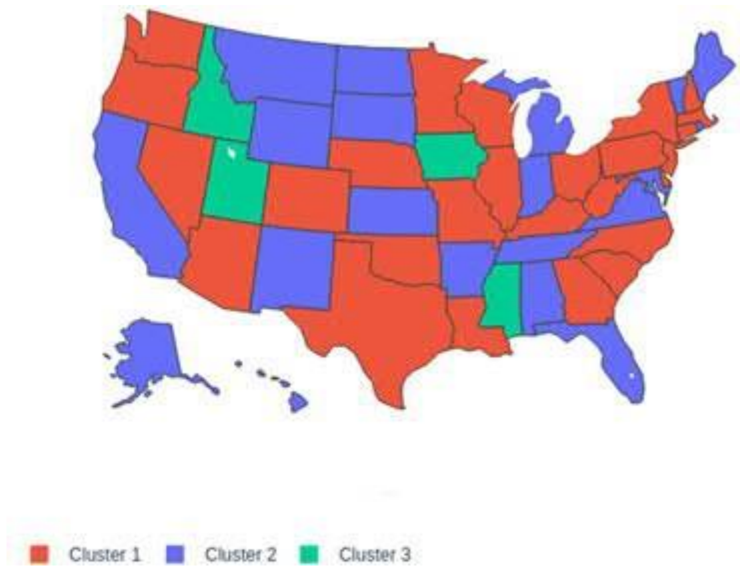
*The red line represents BIC, and the blue line represents AIC

As *Figure 1* demonstrates, there is a significant increase in model fit moving from two to three clusters, and then model fit degrades slightly when additional clusters are added through six clusters. A significant decrease in model fit is observed at seven clusters and an increase from seven to eight, and then a leveling-off is seen between eight and ten. As the plot shows, fit degrades more significantly as additional clusters are added when looking at BIC as compared to AIC. This is expected given that BIC penalizes model complexity more heavily than AIC (Kuha, 2004), so adding additional clusters to the model is more punitive for BIC scores than AIC. Taken collectively, this analysis proceeds utilizing the GMM specification employing three clusters. That specification represents a good balance of model fit and simplicity: It is the overall preferred solution according to BIC across all specifications, and, while not the lowest in absolute score for the AIC, represents a local minimum below the models employing seven or more clusters, numbers that seem unwieldly for explanatory purposes in an analysis looking at only 50 states plus the District of Columbia.

RESULTS

Figure 2 presents the results generated by the three-cluster GMM as a choropleth map, displaying the geographic distribution of the three identified clusters across the United States.

Figure 2: Choropleth Map of the Identified State GMM Clusters



Before engaging in exploration of the composition of the clusters, I offer some commentary about overall model fit of the estimated GMM. An advantage of the GMM procedure is that it is probabilistic, implying that it is possible to inspect the probabilistic assignments of states to clusters. In other words, one can inspect the model's assigned probabilities that a given state belongs to a given cluster. Inspection of these probabilities is important because the algorithm will assign each state to the cluster with the highest assigned probability, regardless of the relative difference between that probability of assignment and the probability of the next best-fitting cluster. As an example, with three clusters, one would want to exercise a high degree of caution in placing great substantive meaning on cluster assignments if, for example, the probabilities of assignment for the observations were roughly 33% to each of the three clusters, with one slightly higher than the other two. While I do not provide the full probability table associated with the final estimated model here. Inspection of it reveals it is safe to proceed with interpreting the results with a high degree of certainty: The lowest observed probability for the assignment of a state to its actual, final cluster is over 99%. In other words, the algorithm was able to separate the individual states into the observed clusters with a high degree of confidence.

As the map conveys, Cluster 1 is the modal category, representing 27 states. Clusters 2 and 3 represent 20 and four states, respectively (the District of Columbia is treated as the 51st state). *Table 2* below presents the results of the GMM clustering exercise, displaying the median levels of stringency for each of the four substantive categories for each of the three identified clusters as initial evidence for the purposes of exploring and evaluating the competing hypotheses described above.

Table 2: Median Stringency Scores by Cluster

Cluster	Organizational & Governance	Academic	Consumer Protection	Student Outcomes
1	15.0	19.5	11.0	3.3
2	12.4	13.6	8.5	0.0
3	8.3	10.0	9.0	2.5

In interpreting *Table 2*, two things are important to keep in mind. First, because the scores for each dimension are scaled differently as they represent a composite of a differing number of indicators, one cannot directly compare a score in one dimension to a score in another. So, for example, one cannot look at the scores for Cluster 1 and say that those states, on average, are nearly five times more stringent with respect to organizational and governance features than they are with respect to student outcomes. Each stringency score can only be interpreted relatively within its own dimension. Second, it is important to remember that Ness, Baser, and

Dean (2021) generate relative, rather than absolute, evaluations of stringency. This means that one can draw upon their data to evaluate a particular state's stringency in a particular dimension as compared to another's and engage in analysis of a state's relative positioning within the identified dimensions, but one cannot extrapolate beyond this to say that the state is stringent relative to any external, absolute standard either within a particular dimension or in general.

With these caveats under consideration, I do observe some interesting patterns in terms of the nature of the classifications generated by the GMM. First, speaking broadly, the Cluster 1 states are, on average, comprised of the states that are the most stringent across all four authorization dimensions, and Clusters 2 and 3 generally represent incremental decreases in stringency across the board. However, looking at the individual dimensions, it becomes apparent that the GMM is relying upon differences in the stringency of the organizational and governance and academic metrics scores as there is relative homogeneity in the scores of the other two dimensions across the clusters. Alignment is also observed within clusters. Cluster 1 states, on average, have higher stringency scores in both operations and governance and academics than Cluster 2 and Cluster 3 states, and Cluster 2 states are higher in both than those in Cluster 3. In terms of scale, Cluster 3 states score roughly half as high as Cluster 1 states in these two dimensions.

Taken collectively, then, some preliminary empirical support is identified for the idea that states approach authorization stringency across substantive dimensions as strategic complements: States that exhibit a high level of stringency (measured both in terms of breadth and intensity of regulatory function) in one dimension tend to exhibit similar proclivities with respect to stringency in others. In other words, states that stringently regulate one area of authorization tend to do so in others. And, while there is a significant degree of heterogeneity in measured stringency across the states and domains, the results lend support to the idea that assessment of stringency levels is meaningful for the purpose of classification.

RESEARCH QUESTION 2

Introduction and Theoretical Framework

Having identified empirical evidence of clustering in the stringency of authorization approaches taken by the individual states, an obvious next step is to ask why some states adopt a more aggressive and stringent regulatory posture than others. Fortunately, a rich literature in higher education, political science, economics, and public policy provides numerous potentially fruitful avenues for undertaking such efforts. For parsimony, I have distilled from this literature four primary categories of relevant characteristics to explore: state political dynamics, state socioeconomic and demographic factors, higher education system characteristics, and state regulatory environments.

That the regulatory actions taken by governments are the product of political processes is well acknowledged, implying that political characteristics are relevant for understanding regulatory outputs and outcomes. Scholarship has recognized at least three unique, nested arenas where political factions fight for influence over the “who” and the “how:” what organizations are to be subjected to regulation, and what mechanisms of control will be employed. First, there is the battle for control of the legislative process itself, where political actors seek to influence the direction and intensity of legislation. Second, there is the space between electorally accountable institutions and the unelected regulators who create and enforce rules. And third, there is the space between regulators and the individual entities subject to regulation (Carrigan & Coglianese, 2011). Accordingly, assessment of the relative differences between the political dynamics of the states in terms of the power dynamics within these nested arenas presents a potentially fruitful avenue for understanding differences in regulatory design and control.

The socioeconomic and demographic characteristics of citizens have been shown to influence policies in several substantive domains, including regulation and redistribution, among others (Alesina & Giuliano, 2011; Kahn 2002). A variety of theoretical pathways connect the two. Basic conceptualizations of the rational actor model suggest that all else equal, individuals will prefer policies that provide benefits to themselves and individuals like themselves, so it stands to reason that the differences in the characteristics of the populations of states would result in differences in policy. Further, a rich vein of literature demonstrates the existence of systematic differences in the policy preferences of different groups, allowing us not just to acknowledge the potential for these differences to percolate up into policy but to begin to systematically explain and understand the nature and direction of these relationships. In particular, states with population concentrations of groups with higher proclivities to patronize institutions of higher education seem likely to support policies that support and protect students.

It seems obvious that other relevant features of state higher education systems may influence the domain of authorization. However, initial efforts at connecting the two have raised some important questions about the coherency and cohesion of authorization within the broader higher education landscape. Ness, Baser, and Dean (2021) document the differences in structural positioning of authorization agencies in state governments, noting that in some states, there appears to be a tight integration with the existing higher education governance infrastructure, and in other states, authorization seems to be wholly independent from those systems. Nonetheless, as pressure for increased access and accountability mounts, states may face pressure to develop informal or formal mechanisms of coordination where such networks do not currently exist. Further, it stands to reason that states with more developed and better-supported systems of higher education may have a systematically different approach to authorization than their lesser-supported peers.

Finally, a competing narrative to that presented in the previous paragraph is that it may be the regulatory function of authorization that defines its dominant characteristic, not its substantive application to higher education. Seen from that perspective, it stands to reason that a state's broader posture toward regulation across domains, such as environment, business, and even primary and secondary education, could itself function as a powerful explanatory factor in understanding the differences in authorization stringency. In other words, authorization may be seen by states as a regulatory policy applied to higher education, rather than a higher education policy dealing with regulation.

Moving forward with evaluating these alternative theoretical perspectives, one important limitation of this analysis is the mismatch between the dynamic nature of policy change and the inherently static process of data collection and analysis. State policies—including, but not limited to, regulatory posture and stringency—generally seem to change incrementally up until the point at which they do not. Point-in-time measures of policies and other characteristics yield only valid measures for the moment at which they were created. Time series and panel data methods utilize repeated observations to make stronger claims about the directionality of the causal arrows between variables than cross-sectional analyses can achieve. The analyses presented here are built upon cross-sectional research designs and draw on individual indicators representing a range of points in time, driven primarily by availability. In general, the selection of these variables represents an attempt to utilize a variety of indicators in order to capture the essence of the breadth and depth of the latent categories described above.

Accordingly, the analysis here, to the extent possible, does not utilize the language of causality, although I will talk about independent and dependent variables when discussing the specification and estimation of the empirical models—but only to indicate which variables appear on the right and left sides of the estimated equations. The use of that language is not intended to suggest that individual variables function as meaningful predictors (or drivers) of the adoption of particular configurations with respect to stringency in authorization. Rather, the individual indicators should be considered collectively as an attempt to represent the latent categories described above. In other words, the empirical exercise which follows does not answer the question, “Does a particular individual variable predict or explain the adoption of a particular configuration of authorization stringency over the others?” Instead, it seeks to explore whether the broad categories discussed above (state political dynamics, state socioeconomic and demographic factors, state higher education system characteristics, and broader state regulatory environments) are related to authorization stringency, thereby taking an initial first step in better defining and explaining the complex underlying dynamics surrounding authorization.

Methods and Model Specification

To this end, I estimate four unique multinomial logistical regressions—one for each category of variables—utilizing a nominal variable capturing the three authorization stringency clusters identified by the GMM above as the dependent variable. If GMMs represent a form of unsupervised machine learning designed to explore undiscovered latent patterns of clustering, multinomial logistical regression can be thought of as a complementary exercise in supervised machine learning, with the goal not of identifying latent clustering patterns, but utilizing observables to build explanatory models predicting qualitative categorizations. To put it differently, the GMM exercise identified the latent groupings in the dataset, and the multinomial logistical regressions explore the factors that are systematically associated with assignment to one cluster over the others across states. Individual regressions are specified for each category of variables in order to explore the unadjusted effect of each category on authorization stringency, avoiding issues of indirect causal pathways and multicollinearity.

As independent variables in the multinomial logistical regression models, I chose for each category three variables designed to capture key features of the underlying categories (parsimony in model specification being important in estimating models that fit three categories measured over 50 observations). *Table 3* displays the individual indicators utilized, providing the variable names, the calendar year from which they were collected, and the source of the measure.

Table 3: Independent Variables

Variable	Year	Source
State political dynamics		
Citizen ideology	2013	Berry, Ringquist, Fording,and Hansen (1998)
Legislative professionalism	2003	Squire (2007)
Education interest groups per capita	2017	Holyoke (2019)
State socioeconomic and demographic factors		
Population age 19-25	2017	U.S. Census Bureau
Per capita income	2017	Bureau of Economic Analysis
State population	2017	U.S. Census Bureau
State higher education characteristics		
Higher education governance centralization	2013	Lacy (2011)
Logged total higher education enrollments	2017	National Center for Education Statistics
State higher education appropriations per capita	2017	SHEEO-SHEF
State regulatory environment		
Regulatory freedom score	2009	Mercatus Center
State private school regulation grade	2006	Friedman Foundation
State regulatory keywords index	2016	Sorens, Muedini, and Ruger (2008)

For the state political dynamics category, I include Berry et al.'s (1998) measure of citizen ideology to account for the fact that ideological conservatism is potentially associated with less stringent approaches to regulation. Legislative professionalism is included to account for potential differences in the "carrying capacity" of legislators to design, oversee, and manage more complex regulatory regimes (Borwn & Greene, 2014; Squire 2007). Education interest groups per capita is included to account for the potential influence of lobbying on regulatory design, although the direction of the effect is ambiguous: On one hand, potential new educational providers may prefer less regulation. The regulatory capture literature in economics suggests that existing educational institutions may prefer more stringent regulation to prevent new competitors from entering the market (Brown, 2021).

In terms of socioeconomic and demographic factors, I include the proportion of the state population aged 19-25 because states with greater college-aged populations may face greater public pressure to regulate higher education to promote greater access and affordability and may afford students greater protections from potential predatory practices. Similarly, states with more affluent citizens may have a greater capacity and willingness to engage in more stringent regulation of higher education since participation in higher education increases with income. Larger states likely have more crowded and diverse landscapes for higher education and may therefore prefer a higher degree of regulation to homogenize and protect quality.

Similarly, the characteristics of the state higher education system may also be related to differences in authorization posture and stringency. Lacy's (2011) measure of higher education governance centralization is included to capture differences in state arrangements for managing and overseeing higher education. This measure diverges from the typical typology of state higher education governance into "advisory board," "coordinating board," and "consolidated governing board," and instead utilizes 19 unique indicators and a Bayesian latent variable model that places all states on a common continuum of centralization. A rich body of work has explored the interrelationships between the structural arrangements of state higher education governance and a variety of relevant policy outcomes, concluding that governance structure matters for a surprisingly broad set of outputs and outcomes, albeit not always through direct, linear pathways (Tandberg, 2013). Greater enrollments, all else equal, may create increased visibility and pressure for attention from policymakers. Higher education appropriations per capita may serve as a proxy measure capturing the relative willingness and ability of citizens and legislatures to support higher education among competing demands for public resources.

The final category is designed to capture the potential relationship between a state's approach to regulation (broadly defined) and the stringency with which it approaches authorization. The regulatory freedom score is a composite measure that captures a variety of measures of economic and social freedoms, including property rights protection, occupational licensure freedoms, and the stringency of tort law, among others. It is intended to capture a state's general posture toward regulated versus more laissez-faire approaches to managing private entities. The state private school regulation grade captures exactly that: the stringency with which states regulate private K-12 educational providers. The state regulatory keywords variable (Sorens, Muedini & Ruger, 2008) is a novel index which reflects the comparative prevalence of regulatory language within state statute, providing an alternative source for data on the general proclivity of state legislatures to engage in regulatory activities versus other legislative endeavors.

As *Table 3* reveals, the modal variable represents data collected in the late 2010s, although a few are collected in earlier years. The source column in the table captures the individuals or organizations responsible for the development and/or dissemination of the variable. In the case of a source that predates the year noted in the year column, the reference in the source column represents the party responsible for the initial development of the measure, which was updated using more recent data. Many of the individual indicators represented in *Table 3* were collected as part of the Correlates of State Policy dataset (Jordan &

Grossmann, 2020). The rest were collected directly from the responsible author, agency, or organization. The final dataset represents a complete cross-section of the 50 states. The only difference between the analytical sample here and that utilized in the earlier clustering exercise is the omission of the District of Columbia due to missing data for most of the utilized variables. Finally, Lacy's (2011) measures of higher education governance centralization only capture the contiguous 48 states, so median imputation was employed to generate data for Hawaii and Alaska.

RESULTS

Since the focus here is not on proving patterns of causality or evaluating the relative strength of individual variables versus others, I do not present tables of the individual estimated coefficients and the corresponding tests of statistical significance. Rather, I take a more general approach, reporting, for each model, the results of the Likelihood-Ratio Chi-square test, and, for the models with statistically significant Chi-square test statistics, a confusion matrix. The Likelihood-Ratio Chi-square test evaluates the null hypothesis that the estimated model contains no non-zero coefficients. In other words, a rejection of it indicates that the specified model tells us something useful for understanding which of the three authorization stringency clusters a particular state belongs to. As noted in the literature, evaluation of model fit is not as straightforward as logistic or OLS regression when evaluating the multinomial logit model. Accordingly, the confusion matrix provides a mechanism to evaluate the relative strength of the association between the independent variables in the model and the dependent variable by gauging the extent to which the model can correctly predict the assignment of states to the identified authorization stringency clusters, allowing us to assess its ability to differentiate between the members of the individual clusters based on the specification of the model. *Table 4* presents the Likelihood-Ratio Chi-square tests, and *Tables 5* and *6* display the confusion matrices for the statistically significant categories of variables.

Table 4: Likelihood Ratio Chi-square Test Results

Category	LR Chi-square	Prob > Chi-square
State political dynamics	9.15	0.1654
State socioeconomic and demographics characteristics	17.97	0.0063
State higher education characteristics	21.21	0.0017
State regulatory environment	10.52	0.1043

Table 5: Confusion Matrix, State Socioeconomic and Demographic Characteristics Model

Actual Cluster / Predicted Cluster	Cluster 1	Cluster 2	Cluster 3
Cluster 1	21	10	0
Cluster 2	5	9	2
Cluster 3	0	1	2

Table 6: Confusion Matrix, State Higher Education Characteristics Model

Actual Cluster / Predicted Cluster	Cluster 1	Cluster 2	Cluster 3
Cluster 1	22	9	4
Cluster 2	3	11	0
Cluster 3	1	0	0

As shown in *Table 4*, two of the four identified categories yield a statistically significant Likelihood Chi-square test, indicating a model fit superior to a null model (a model with no covariates): the state socioeconomic and demographic characteristics, and the state higher education characteristics. No significance is achieved for the models that attempt to capture the state's political and regulatory environments. Substantively, this provides some preliminary, suggestive evidence that there is an underlying connection between the characteristics of the population and the characteristics of the higher education system and the state's position with respect to higher education authorization stringency. Looking at the confusion matrices, I calculate the accuracy of each specified model as a back-of-the-envelope mechanism for assessing the strength of the association between each category and authorization stringency. In so doing, summing the observations assigned to the main diagonal of the tables reveals that both models assign 32 of the 50 states to the correct stringency cluster. To provide context, a naïve model would simply assign each observation to the modal cluster, yielding correct classification of 27 states, so our estimated models yield on average an 18.5% improvement in classification compared to that baseline.

The accuracy of prediction can also be calculated individually for each of the three clusters. The socioeconomic and demographic model correctly classifies 80.8% of Cluster 1 states, 45% of Cluster 2 states, and 50% of Cluster 3 states. The state higher education characteristics model correctly classifies 84.6% of Cluster 1 states, 55% of Cluster 2 states and fails to correctly classify any of the Cluster 3 states. In other words, the identified characteristics seem more useful in differentiating between states choosing to adopt a high level of stringency from states choosing a medium or low level of stringency and less so for identifying states choosing medium or low levels versus other states.

DISCUSSION, CONCLUSIONS, AND IMPLICATIONS FOR FUTURE RESEARCH

As is the case with all exploratory research, these findings—distinct clusters in state authorization exist, and there is some association between state characteristics and the assignment of these clusters—raise more questions than they answer. First, of course, is the question of whether these results would be confirmed by additional analysis. Ness, Baser, and Dean (2021) suggest that categorization of states based on the aggregate level of authorization stringency across the identified domains provides a useful way to differentiate states. This research, utilizing an unsupervised machine learning clustering algorithm, provides independent support for that supposition. In general, states with high levels of stringency in one substantive domain tend to also have higher levels in the others, suggesting states may possess a preferred level of regulatory oversight in authorization that is then consistently applied across domains. However, additional work is needed to test this hypothesis. If the limited extant research on state authorization demonstrates anything, it shows a surprising breadth of approaches and arrangements to authorization across the states, a fact seemingly counter to the rational, purposeful application of regulatory authority suggested here. Whether the observed clustering is the result of intentional design or accidental consequence is certainly worthy of additional scrutiny, and the data collected here are unable to differentiate between the two. Complementing existing work with further qualitative analyses will allow us to develop a deeper understanding of not just the observable features of state authorization approaches, but also to develop a deeper understanding of the various pathways that lead states to adopt particular authorization regimes as opposed to others.

Turning to the second research question, it is also interesting and perhaps counter to expectation that significant associations are not identified between authorization stringency and state political and regulatory characteristics. Recent events have reinforced the intensity of the political dynamics surrounding higher education in general, but perhaps, consistent with the lack of scholarly attention paid to authorization as a substantive domain of pragmatic importance, authorization similarly exists outside the scope (or below the radar) of political actors. Similarly, it is surprising that the broader state regulatory climate was not found to have a significant association with regulatory stringency. Higher education policy scholars have long argued that higher education receives differential treatment as compared to other policy domains by state governments, and perhaps this finding is further confirmation of the scope of these differences. Alternatively, it could be that the research design and particular indicators chosen simply fail to capture the underlying political dynamics at play within a state. The utilization of panel data that capture changes in state dynamics and changes in authorization approach would allow us to better connect the two by taking advantage of the dimension of time to infer stronger causal relationships. The additional degrees of freedom associated with this larger dataset would also permit the inclusion of additional indicators and allow us to specify more robust models that test not only the strength of relationships holding other features constant, but also allow the possibility for interactions between categories.

Lastly, this research ultimately can say little to answer the question of best practice. Theoretically, one might suppose that a high degree of stringency across the individual domains should yield a tighter coupling between the public policy preferences underpinning authorization and the behaviors of the regulated institutions of higher education. However, regulations impose numerous costs that accrue both to the regulatory body as well as the regulated organization. In other words, when it comes to regulatory stringency, it is not automatically the case of more equaling “better,” at least when better is described in terms of efficiency and cost-effectiveness. Further, the reality of the complex nature of higher education implies that even the most stringent regulatory instruments will be imperfect, presenting an opportunity for regulated institutions to recoup the additional costs imposed by regulation by reducing quality in ways that escape the reach of regulation. The relationship between regulatory stringency in authorization and quality of higher education—variously defined—represents an area ripe for future research.

REFERENCES

Alesina, A., & Giuliano, P. (2011). Preferences for redistribution. The Handbook of socialeconomics (Vol. 1, pp. 93-131). North-Holland.

Berry, W. D., Ringquist, E. J., Fording, R. C., & Hanson, R. L. (1998). Measuring citizen and government ideology in the American states, 1960-93. *American Journal of Political Science*, 327-348.

Bowen, D. C., & Greene, Z. (2014). Should we measure professionalism with an index? A noteon theory and practice in state legislative professionalism research. *State Politics & Policy Quarterly, 14*(3), 277-296.

Brunel, C., & Levinson, A. (2013). Measuring environmental regulatory stringency. OECD Trade and Environment Working Papers, 2013/05, OECD Publishing.

Brown, J. (2020). A Blind Eye: How the Rational Basis Test Incentivizes Regulatory Capture in Occupational Licensing*. Journal of Law, Economics and Policy*, Forthcoming.

Burnham, K. P., & Anderson, D. R. (2004). Multimodel inference: understanding AIC and BIC in model selection. *Sociological methods & research, 33*(2), 261-304.

Carrigan, C., & Coglianese, C. (2011). The politics of regulation: From new institutionalism tonew governance. *Annual Review of Political Science, 14*, 107-129.

Cohen, M. D., March, J. G., & Olsen, J. P. (1972). A garbage can model of organizationalchoice. *Administrative science quarterly*, 1-25.

Grossback, L. J., Nicholson-Crotty, S., & Peterson, D. A. (2004). Ideology and learning in policy diffusion. *American Politics Research, 32*(5), 521-545.

Holyoke, T. T. (2019). Dynamic state interest group systems: A new look with new data.

InterestGroups & Advocacy, 8(4), 499-518.

Jordan, M. P., & Grossmann, M. (2020). The Correlates of State Policy Project v.2.2. East Lansing, MI: Institute for Public Policy and Social Research (IPPSR).

Judge-Lord, D., McDermott, C. L., & Cashore, B. (2020). Do private regulations ratchet up?How to distinguish types of regulatory stringency and patterns of change. *Organization & Environment, 33*(1), 96-125.

Kahn, M. E. (2002). Demographic change and the demand for environmental regulation. *Journalof Policy Analysis and Management, 21*(1), 45-62.

Kelly, A., James, K., & Columbus, R. (2015). *Inputs, outcomes, quality assurance: A closer look at state oversight of higher education.* Washington, DC: American Enterprise Institute.

Kuha, J. (2004). AIC and BIC: Comparisons of assumptions and performance. *Sociologicalmethods & research*, 33(2), 188-229.

Lacy Jr, T. A. (2011). *Measuring state postsecondary governance: Developing a new continuumof centralization* (Doctoral dissertation), University of Georgia.

McCubbins, M. D., & Schwartz, T. (1984). Congressional oversight overlooked: Police patrolsversus fire alarms. *American Journal of Political Science*, 165-179.

McLachlan, G. J., & Basford, K. E. (1988). Mixture models: Inference and applications toclustering (Vol. 38). New York: M. Dekker.

McLendon M.K. (2003) State Governance Reform of Higher Education: Patterns, Trends, and Theories of the Public Policy Process. In: Smart J.C. (ed). *Higher Education: Handbook of Theory and Research*, vol 18. Springer, Dordrecht.

https://doi.org/10.1007/978-94-010-0137-3_2

McLendon, M. K. (2003). Setting the governmental agenda for state decentralization of highereducation. *The Journal of Higher Education, 74*(5), 479-515.

Ness, E., Baser, S. & Dean, M., (2021). *State authorization landscape and process: An inventory, classification, and analysis*. Boulder, CO: State Higher Education Executive OfficersAssociation.

Shipan, C. R., & Volden, C. (2008). The mechanisms of policy diffusion. American Journal of Political Science, 52(4), 840-857.

Sorens, J., Muedini, F., & Ruger, W. P. (2008). US state and local public policies in 2006: A new database. *State Politics & Policy Quarterly, 8*(3), 309-326.

Squire, P. (2007). Measuring state legislative professionalism: The squire index revisited. *StatePolitics & Policy Quarterly, 7*(2), 211-227.

Tandberg, D. A. (2013). The conditioning role of state higher education governance structures. *The Journal of Higher Education, 84*(4), 506-543.

Tandberg, D. Brueckner, E., & Weeden, D. (2019). Improving state authorization: The state role in ensuring quality and consumer protection in higher education. Boulder, CO: StateHigher Education Executive Officers Association.

Teixeira, P. N. (2006). Markets in higher education: can we still learn from economics' foundingfathers?. Berkeley, CA: Center for Studies in Higher Education (CSHE).

Thelin, J. R. (2011). *A history of American higher education*. JHU Press.

Toutkoushian, R. K., & Shafiq, M. N. (2010). A conceptual analysis of state support for highereducation: Appropriations versus need-based financial aid. *Research in*

Higher Education, 51(1), 40.

Volkwein, J. F., & Tandberg, D. A. (2008). Measuring up: Examining the connections among state structural characteristics, regulatory practices, and performance. *Research in HigherEducation, 49*(2), 180-197.