DOI: https://doi.org/10.53555/nncse.v9i5.1696 Publication URL: https://nnpub.org/index.php/CSE/article/view/1696

STATISTICAL METHOD FOR EMPIRICAL TESTING OF COMPETING THEORY IN ANAMBRA STATE POLYTECHNIC, MGBAKWU.

Okoli, Boniface Chukwuma[1*], Enukonwu Patience Amaoge[2]

[1,2]Institution Anambra State Polytechnic, Mgbakwu, Dapartment Computer Engineering Technology,*
Telephone No: 0813128493, E-Mail : okolibonifacechukwuma@gmail.com

Corresponding Author:
okolibonifacechukwuma@gmail.com

Abstract

Empirical testing of competing theories lies at the heart of social science and applied Science research. We demonstrate that a well-known class of statistical models, called finite mixture models, provides an effective way of rival theory testing. In the proposed framework, each observation is assumed to be generated either from statistical models implied by one of the competing theories or more generally from a weighted combination of multiple statistical models under consideration. Researchers can then estimate the probability that a specific observation is consistent with each rival theory. By modeling this probability with covariates, one can also explore the condition under which a particular theory applies. We discuss a principles way to identify a list of observations that are statistically consistent with each theory and propose measure of the overall performance of each competing theory. We illustrate the relatives' advantages of our method over existing methods through empirical and stimulation studies.

Keywords: *Empirical testing; Competing theories, statistical method; finite Mixture-Models; probability model.*

INTRODUCTION

mpirical testing of competing theories lies at the heart of social science research. Since there typically exist alternative theories explaining the same phenomena, researchers can often increase the plausibility of their theory by empirically demonstrating its superior explanatory power over rival theories. In political science, Clarke set forth this argument most forcefully by claiming that “theory confirmation is not possible when a theory is tested in isolation, regardless of the statistical approach” (2007b, 886). In order to quantitatively test competing theories, however, most political scientists fit a regression model with many explanatory variables that are derived from multiple theories. Achen (2005) strongly condemns such practice as atheoretical and calls it a “garbage-can regression.” To address this critique, some researchers have relied upon various model comparison procedures to assess the relative performance of statistical models implied by different theories under investigation.

Problem Statement/Justification

Over the past decade, peoplple have been questioned the uses of statistiscal testing in sciences.the criticisms apply to both experimental data and treatments, random assignment of experimental units, replication. The most glaring problem with the of hypothesis testing is that nearly all null hypothesis are obviously false on a priori ground

Objective(s) of the study

Aims and Objectives

A Statistical Method for Empirical Testing of Competing Theories

S*pecifics Objectives*

- to review the specification, estimation, and inference for finite mixture models in the context of empirical testing of competing theories.
- to discuss a method to identify the observations that are statistically significantly consistent with each theory.
- We also propose several ways to measure the overall performance of each competing theory.
- Finally, we compared the proposed approach with the standard model selection procedures.

Literature Review:

In this section, we briefly describe the background of the motivating empirical example regarding the competing theories of trade policy preferences. An enduring theme in the international political economy literature is the explanation of preferences for free trade. In a seminal contribution, Hiscox (2002) analyzes legislative voting on trade bills in the United States by drawing on political economy interpretations of two canonical theories from the trade literature: the Stolper-Samuelson (SS) and Ricardo-Viner (RV) models of international trade. The two competing theories differ critically in the extent to which they emphasize *factoral* versus *sectoral* cleavages. The SS model suggests that cleavages on trade policy will be along factoral lines and predicts that the owners of factors which the United States is relatively abundant in (compared to the rest of the world) will favor trade liberalization.[6] In contrast, the RV model suggests an alternative cleavage between supporters and opponents of free trade that runs along sectoral lines. These two models of support for trade policy figure centrally in this long tradition of international political economy research (e.g., Ladewig 2006; Rogowski 1989; Scheve and Slaughter 2001).

Gordon and Smith (2004), emphasize that the proposed approach does not require researchers to specify such conditions. In fact, researchers may wish to use the proposed mixture modeling approach in order to explore what factors determine the relative applicability of each rival theory. Of course, the findings obtained from such exploratory analyses must be interpreted with caution, and deductive and systematic theory testing is required to draw more definitive conclusions Those in export industries should favor liberalization, whereas those in import competing A key observation made by Hiscox (2002) is that the applicability of these competing models depends on how *specific* factors of production are to particular industries. If capital is highly mobile in the national economy, meaning it can easily move across industries, then the SS model is likely to be supported because the winners and losers of trade will be found among owners of abundant and scarce forms of factors, respectively. On the other hand, if capital is more specific (i.e., less mobile), then cleavages should fall along sectorial lines since capital is unable to easily adjust across industries. Hence, Hiscox hypothesized that whether congressional voting on trade bills can be explained by the SS or RV model will depend on the degree of factor specificity in the U.S. economy

Methodology

In this section, we first briefly review the specification, estimation, and inference for finite mixture models in the context of empirical testing of competing theories. Before describing the proposed methodology, we emphasize an important distinction between *causal* and *predictive* inferences. For causal inferences, ignoring relevant confounders may result in omitted variable bias. In contrast, the existence of omitted variables alone does not invalidate predictive inferences. Indeed, it is well known.

*‘‘*Although controlling for irrelevant variables can sometimes result in bias, in typical observational studies analyzed by social scientists it is difficult to argue that researchers should not adjust for the observed differences between the treatment and control groups; suppose that we have a simple random sample from a population and find that in this sample, the turnout of overweight voters is significantly lower than that of other voters. It clearly known that for the purpose of *predictive* inferences, parsimonious models tend to outperform unnecessarily large models (see, e.g., Hastie,

Tibshirani, and Friedman 2001). Thus, if the goal of researchers is to construct a theory with strong predictive power (as opposed to testing causal mechanisms; see Imai, Tingley, and Yamamoto 2011, for relevant methodological issues), parsimonious models that can capture systematic patterns in the data are preferred. While our method can be used for both purposes, the causal inference approach would require strong research designs that enable the identification of causal effects. Our empirical examples should be thought of as the instances of predictive inference. Nevertheless, whenever using mixture models, well-specified theories play an essential role in model specification.

Finite Mixture Models:

Model Specification**.** Consider a finite number of M different statistical models, each of which is implied by one of the competing theories explaining the same phenomena. Beyond the fact that it can handle more than two theories at the same time, the proposed method is applicable without modification regardless of whether these statistical models are nested or not Finite mixture models are based on the assumption that each observation is generated either from one of the M statistical models or more generally from a weighted combination of multiple statistical models. This does not necessarily imply that researchers must identify all relevant theories. It is also possible that any observation, which is consistent with one of M theories under consideration, is also consistent with other theories that may or may not be included in the analysis. Rather, the goal of finite mixture models is to measure the *relative* explanatory power of the competing theories under consideration by examining how well a statistical model implied by any of the rival theories predicts each observation in the sample. For example, it is perhaps the case that the Stolper-Samuelson and Ricardo-Viner theories do not exhaust all possible theories for trade policies. And yet, it is of interest to investigate the relative performance of each theory explaining the variation in the voting behavior of legislators.

Formally, let $f_m(y \mid x, \theta_m)$ denote a statistical model implied by theory m where y is the value of the outcome variable Y , x is the value the vector of covariates X takes, and θ_m is the vector of model parameters. In statistics, typical applications of finite mixture models involve the same distributional and functional-form assumptions with identical covariates. Similar to random coefficient models, such an approach makes parametric models flexible by allowing different groups of observations to have different parameter values. However, these alternative statistical methods can neither provide a measure of overall support for each theory nor classify each observation to one of the competing theories. Furthermore, different theories usually require different sets of predictors. In fact, Hiscox (2002) employed logistic regression models with different sets of covariates for the Stolper-Samuelson and Ricardo-Viner theories. One may also wish to specify different statistical models for rival theories. For example, when analyzing the duration of cabinet dissolution, the underlying risk of cabinet dissolution may be constant (as in the exponential model)

In this reasearch, we consider a mixture of regressions, which is also known as switching regressions (e.g., Quandt 1972). In physcal science, Brandt and Freeman (2006) use Markov switching model (particular mixture models) for time-series data. To empirically test this hypothesis, Hiscox collected the data on factor specificity in the U.S. economy over nearly two centuries. His measures varied considerably over time, suggesting that during some eras voting should be along factor lines (capital/land versus labor) and in other eras along sectoral lines (exporters versus importers). To leverage these changes over time, Hiscox estimated separate regressions for different eras in time. Using a conventional model selection procedure called the J test, Hiscox provides evidence that support for liberalization is best accounted for by the SS model during eras where specificity was low. In contrast, he finds that in periods where specificity was high, the RV is the preferred model.

Although breaking up the votes into different eras constitutes one informal way to test the factor specificity argument, the continuous measure of the factor specificity variable created by Hiscox does not provide natural breakpoints which can then be used to group votes. Thus, any grouping might be criticized as arbitrary. As we demonstrate, finite mixture models offer a relatively straightforward and yet formal way to directly incorporate the factor specificity measure. In particular, mixture models use the level of factorspecificity to predictwhether the SS or RV model is appropriate for each trade bill or even each vote. Thus, in addition to the overall assessment of the two models, we are also able to identify the list of trade bills in which the voting pattern is consistent with each theory.

Notable exceptions include the work by Hill and Kriesi (2001), Kedar (2005), and Iaryczower and Shum (200). Stated that in *the United States,* capital and land owners will support free trade, whereas those specializing in labor should oppose liberalization

Thus, Z_i can take one of M values, i.e., $Z_i \in \{1, 2,..., M\}$, depending on which statistical model generates the i th observation. The data-generating process is given by

$$Y_i \mid X_i , Z_i \sim f_{Z}i\ (Y_i \mid X_i , \theta_{Z}i \qquad \text{.................................(1)}$$

for each $i = 1,..., N$.

Next, assuming the conditional independence across observations given the covariates and the latent variable, the model specified in equation (1) yields the following observed-data likelihood function where the latent variable Z_i has been integrated out,

L_{obs} 1, $\Pi|\{X_i , Y_i\}N \quad {}_{i=1}$

$$\Pr(Z_i = m \mid W_i) = \pi_m(W_i , \dagger_m), \qquad \text{………………………………………………………………..(2)}$$

Estimation and Inference

Estimation and can proceed by either a frequentist approach of maximizing the observed-data likelihood function or a Bayesian approach of sampling from the posterior distribution after specifying prior distributions. To obtain the maximum likelihood estimates, the Expectation-Maximization (EM) algorithm (Dempster, Laird, and Rubin 1997), an iterative numerical optimization algorithm consisting of the expectation (or E) step and the maximization (or M) step, can be applied to the following complete-data log-likelihood function, which is derived

1. Sample Z_i given the current values of all parameters with the following probability,

Pr $Z(t) = m \mid 1(t-1), \Pi(t-1), \{Yi , Xi\}N =$

Grouped Observations. In some situations, multiple observations are grouped and researchers may wish to assume that all observations of one group arise from the same statistical model implied by either a particular theory or more generally from the same weighted combination of statistical models under investigation. For example, multiple observations may be collected over time for each individual in a study, and all observations from one individual are assumed to be consistent with one of the competing theories or a particular weighted combination. of these theories. In the trade policy example discussed In such a situation, finite mixture models can be formulated as,

$$Y_{ij} \mid X_{ij} , Z_i \sim f_{Zi} (Y_{ij} \mid X_{ij} , \theta_{Zi}), \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \quad (3)$$

for $i = 1,\ldots, N$ and $j = 1,\ldots, J_i$ where

Result(Expected outputs/Results

Identification of Observations Consistent with Each Theory

One advantage of the proposed mixture modeling approach is its ability to yield a list of observations for which researchers have sufficiently strong evidence that they are consistent with one of the competing theories. To do this, we focus on the posterior probability, $_{i,m}$, that observation i is consistent with theory m. This parameter can be estimated as part of either the EM algorithm or the MCMC algorithm in equation (i).(ii). & (iii).

Estimated Population Proportion of Observations Consistent with Model 1 (four left plots) and Classification Success Rates (four right plots) in the Two-Theory Mixture

Identification of Observations Consistent with Each Theory

One advantage of the proposed mixture modeling approach is its ability to yield a list of observations for which researchers have sufficiently strong evidence that they are consistent with one of the competing theories. To do this, we focus on the posterior probability, $_{i,m}$, that observation i is consistent with theory m. This parameter can be estimated as part of either the EM algorithm or the MCMC algorithm in equation (i).(ii). & (iii).

Estimated Population Proportion of Observations Consistent with Model 1 (four left plots) and Classification Success Rates (four right plots) in the Two-Theory Mixture

Estimated Population Proportion of Observations Consistent with Model 1 (two left plots) and Classification Success Rates (two right plots) without a Theory-Predicting Variable in the Three-Theory Mixture Model Simulation Study

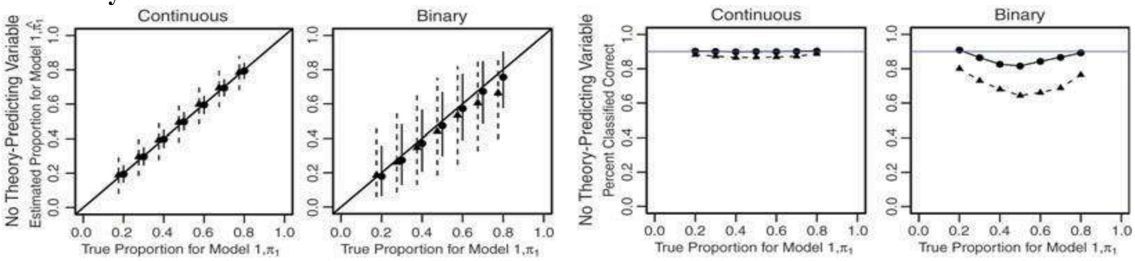
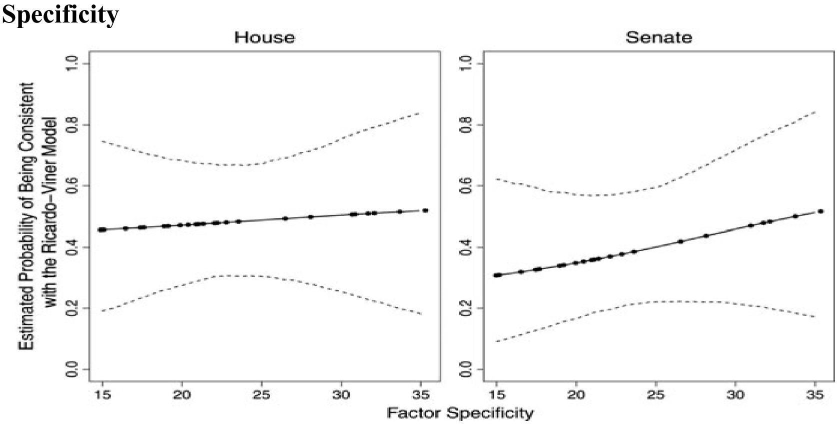


FIGURE 3 Estimated Probability of Votes for a Bill Being Consistent with the Ricardo-Viner Model as a Function of Factor Specificity



Note: Solid line is the estimated probability with actual observations indicated by sold circles, and dashed lines represent 95% confidence intervals based on the Monte Carlo approximation. Although there is a considerable degree of uncertainty due

Comparison with Other Methods.

Finally, the mixture modeling approach also yields estimated model parameters for each of the competing theories, i.e., θ_{SS}, θ_{RV}, as well as the estimated coefficients on variables that are used to estimate mixing probabilities, i.e., $\dagger_{RV}$. In Table 1, we report these estimates and compare them to the "garbage-can" regression (the last four columns), which Achen (2005) and others (e.g., Clarke 2000; Gordon and Smith 2004) argue should be avoided. Here, the "garbage-can" regression refers to the single logistic regression, which contains all five variables taken from both SS and RV models. Following Hiscox's original analysis, we also include bill fixed effects in this model. The table shows that for the mixture modeling approach, *all* estimated coefficients of the two models have expected signs and are statistically significant. For example, the estimated coefficient for the farm variable is negative, implying that states with high levels of agricultural production are more likely to oppose protectionism as expected under the Stolper-Samuelson model. In contrast, in the "garbage-can" regression the coefficients are considerably smaller and their standard errors are larger (relative to the size of the coefficients). For example, the farm variable is not statistically significantly different

from zero both in the House and Senate. This suggests the superior discrimination power of each variable in the mixture model despite the fact that the "garbage-can" regression was fit to the entire data.

The results based on the mixture model also improve upon those reported in the original article. For example, the application of the *J* test indicates that the SS model is selected for the period between 1945 and 1962. However, Hiscox found that the farm and manufacture variables in this model have opposite signs than what is predicted (2002, 603). In contrast, the results of the mixture model show no such inconsistency. Furthermore, when the SS model (with bill fixed effects) is fitted to the subset of votes classified to the RV model given in the second and fourth columns of Table 1 of the supporting materials, the estimated coefficient for the farm variable has a positive sign (statistically insignificant in the House and

`We also ran the same "garbage-can" regression model with the interaction terms between the factor variable and each of the covariates. The results are somewhat puzzling. The coefficients of some main effects do not have expected signs, and others are no longer statistically significantly different from zero. In addition, some signs of the coefficients for these interaction terms are not in the expected direction

TABLE 1 Parameter Estimates and Their Standard Errors from the Mixture Model for the House and Senate

Statistical Analysis.

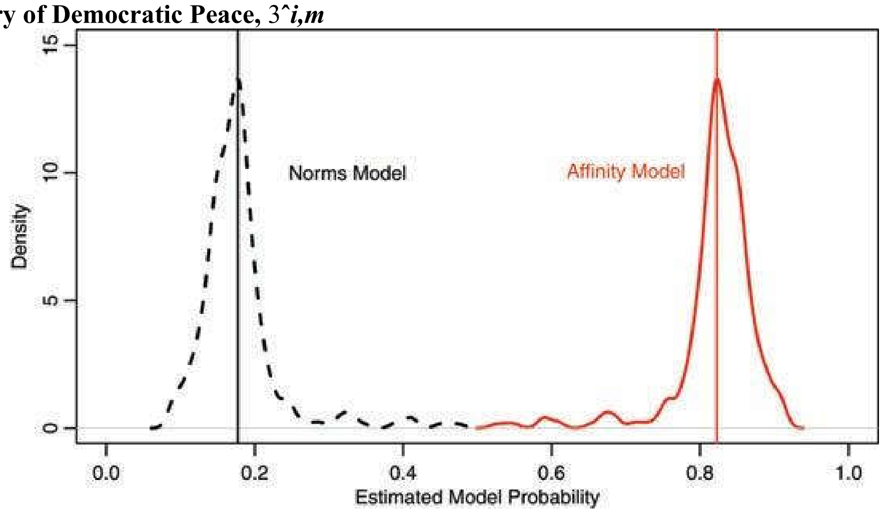
Based upon the lack of statistical significance of the key estimated coefficient and its wrong sign, Huth and Allee conclude that out of the three models the affinity model "produced the weakest results" (2002, 283) and that between the accountability

		Mixture Model				"Garbage-can" Models			
		House		Senate		House		Senate	
Variables		coef.	s.e.	coef.	s.e.	coef.	s.e.	coef.	s.e.
Stolper-Samuelson	intercept	−0.23	0.14	0.02	0.21	0.47	0.12	0.78	0.25
	profit	−1.60	0.53	−5.69	1.19	−0.93	0.56	−3.58	1.23
	manufacture	17.60	1.54	19.79	2.59	10.01	1.11	7.82	2.27
	farm	−1.33	0.29	−1.27	0.43	−0.14	0.24	−0.03	0.42
Ricardo-Viner	intercept	−0.61	0.05	−0.83	0.13				
	import	3.09	0.33	2.53	0.80	1.03	0.34	2.22	0.76
	export	−0.85	0.16	−2.80	0.77	−1.45	0.14	−2.58	0.36
Mixture Probability	intercept	−0.39	1.48	−1.60	1.62				
	factor	0.01	0.06	0.05	0.07				

Statistical Analysis.

Based upon the lack of statistical significance of the key estimated coefficient and its wrong sign, Huth and Allee conclude that out of the three models the affinity model "produced the weakest results" (2002, 283) and that between the accountability.

FIGURE 4 Smoothed Histograms of Estimated Probabilities That Each Observation Is Consistent with Each Competing Theory of Democratic Peace, 3ˆ*i,m*



Note: Solid vertical lines represent the estimated overall probability that observations are consistent with each model, $\hat{\pi}_m$. The affinity model receives the greatest support. The estimated probability for the accountability model is essentially zero for all observations.

In this section, First, the proposed mixture modeling approach provides one way to assess the relative *predictive* performance of rival theories, but like any statistical method, the method in itself does not solve endogeneity and other fundamental problems of *causal* inference in observational studies. For example, one may estimate causal effects using a mixture model which consists of causal submodel Second, one should not test too many competing theories at once. Fitting a mixture model demands much more from the data than fitting each of the submodels separately. The fact that each submodel is identified does not necessarily imply a mixture of all submodels is identified.

References:

[1]. Achen, C. H.(2005.) “Let’s Put Garbage-Can Regressions and Garbage-Can Probits Where They Belong.”Conflict Management and Peace Science (p): 327–339.
[2]. Beck, N., and J. N. Katz. (2007). “Random Coefficient Models for Time-Series-Cross-Section
[3]. Data:Monte Carlo Experi-ments.” Political Analysis 15(p): 182–195.
[4]. Benjamini, Y., and Y. Hochberg. (1995). “Controlling the False Discovery Rate: A Practical and Powerful
[5]. Approach to Multiple Testing.” Journal of the Royal Statistical Society, Series B 57(p): 289–300
[6]. Braumoeller,B.(2003.) “Causal Complexity and the Study of Politics.” Political Analysis 11(p): 209–333.
[7]. Clarke, K. A.(2000.) “Testing Non-nested Models of International Relations: Reevaluating Realism.”American Journal of Political Science 45(3): 724–44.
[8]. Davidson, R., and J. G. MacKinnon .(1981). “Several Tests for Model Specification in the Presence of Alternative Hypothe-ses.” Econometrica 49(3): 781–93.
[9]. Dempster, A. P., N. M. Laird, and D. B. Rubin.(2017) “Maximum Likelihood from Incomplete Data via the EM Algorithm (with Discussion).” Journal of the Royal Statistical Society, Series B, Methodological p. 1–37.
[10]. Friedman, M. (1966.) “The Methodology of Positive Eco-nomics.“ In Essays in Positive Economics. Chicago: University of Chicago Press,P. 3–160.
[11]. Fru¨hwirth-Schnatter, S.(2007). Finite Mixture and Markov Switching Models. New York: Springer.
[12]. Harrison, G. W., and E. E. Rutstro¨m. (2009.) “Expected Utility Theory and Prospect Theory: One
[13]. Wedding and a Decent Funeral.” Experimental Economics (P): 133–158