Using Empirical Data to Test Theoretically Grounded Operationalizations of Protest in an Agent-Based Model[1]

Brandon Alcorn
Lustick Consulting, Inc.
balcorn@lustickconsulting.com

Miguel Garces
Lustick Consulting, Inc.
miguelag@lustickconsulting.com

Dr. Ian S. Lustick
University of Pennsylvania
Lustick Consulting, Inc.
ilustick@lustickconsulting.com

1 Introduction

Les Miserables–the novel, the musical, and now the movie–has aroused readers and audiences to appreciate how the mass experience of injustice is translated into mobilization for change and why success of such mobilization is so problematic. Appearing in Victor Hugo's classic, and in its contemporary revivals, are all the elements with which students of collective behavior generally, and of protest and civil revolt more specifically, must contend. Whether successful or not, large-scale episodes of collective resistance to established authority structures rely on substantial strain on resources or in the way they are distributed. Depending on how individuals and groups interpret that strain, widespread motivations can develop for making sacrifices or taking risks to change the world—not just one's personal circumstances. To be sure, as the "supply-siders" argue (Collier and Hoeffler, 2004), mass mobilization can also be fueled by ambitions for power and profit, rather than demands for justice, but regardless of the specific motivations of elites, they will virtually always use the language of justice denied to reach the masses they wish to mobilize.

Beyond the rationale for anti-regime mobilization provided by what Neil Smelser called "strain" (Smelser, 1962) and Victor Hugo called injustice, a host of other factors combine to determine the occurrence and success or failure of these episodes. Following Smelser and taking into account the key insights of collective action theory, we are led to ask a number of distinct but related questions:

1. Is there strain in the society pushing some actors toward dissatisfaction with prevailing patterns?

2. Are contextual social, cultural, and economic contexts conducive to bottom-up mobilization?

3. Are there organizational frameworks within which coordination can be achieved?

4. Do norms and heuristics exist to guide individuals to frame risk-taking in the context of particular forms of collective action as appropriate responses to emotions of resentment, outrage, or anger?

5. Are there efficacious precipitating events or signals for the beginning of the mobilization?

[1]Support for this paper was received from the Office of Naval Research through the W-ICEWS program (#N00014-12-C-0066) via Lockheed Martin, ATL as well as the V-SAFT program (#N00014-12-C-0042). Researchers interested in replicating findings should contact the authors for access to the necessary platform, templates, and protocols.

6. Are there cracks or weaknesses in the repressive apparatus or the "agencies of social control?"

7. Can the problem of shirking by hopeful but scared and "rational" potential participants be sufficiently overcome to produce sustained mobilization?

No study of an episode, or even a class of episodes, of collective mobilization can attend systematically to the variables to which each of these questions directs attention. On the other hand, a detailed narrative of any one episode would inevitably deal in some way with each kind of factor. That is one reason for the emotional and political range of *Les Miserables.*

2 Our Models

Each dynamic run of one of our agent-based country models yields a distinctive chronicle, if not an actual narrative. The current set of countries we've modeled for this experiment include Bangladesh, Egypt, Pakistan, the Philippines, Thailand, and Venezuela. These have been designed and deployed to forecast and analyze patterns of different kinds of political mobilization (lobbying, protest, violent attacks). Each future (i.e. each "run") of each model (beginning with the present as t=0 represented by randomized configurations of the best data available by internal administrative unit) details a trajectory of how its particular initialization, combined with sequences of randomized exogenous shocks arising from below the analytic horizon, moves through the immense space of possible states attainable by the model. For each country, we are able to create distributions of hundreds or thousands of these simulated trajectories and analyze them collectively to identify mobilization patterns, likely outcomes, event chains, and other outcomes of interest.

Theories used to build our models combined with data used to populate them provide the warrant for treating the portion of the state space any one model explores as a surrogate for the possible futures of the target country. Standard model inputs include data on ethnic and religious groups, results from recent elections, poverty data, survey data relating to issues of nationalism, support for the state, support for the military, corruption, globalizing tendencies, and business affiliations, and answers to questionnaires by subject matter experts detailing the networks of influence between political elites, the group affiliations associated with the elites, important anti-government groups, and the varied political salience levels of key groups. Whenever possible, these data are geographically specified to ensure the spatial integrity of model results.

The theoretical underpinnings of our country models along with the challenges of operationalizing theories that may be commonly accepted but are often vaguely specified have been discussed in previous publications (Alcorn et al., 2012; Lustick et al., 2012). For example, our models are built upon widely accepted theories within constructivism that individuals have a range of identities that change in salience over time, that current identities can be discarded, and that new identities can be attained.(Chandra, 2012) However, there is very little written and no consensus about how exactly identity changes occur. We know that identity salience can change, but how much environmental pressure is needed for the salience of a current identity to change? How hard is it to attain a new identity or discard an old one? How often do individuals change their identities? These are all questions that require some type of answer in order for constructivist identity theory to be implemented within a formal model. We have answered these questions by estimating parameter values associated with each theory through a combination of testing and evolutionary processes over a decade of model building and experimentation. We have changed parameter values when previous values have proven unsatisfactory. So while there is a large range of well-established theoretical work that informs our models, the models are ultimately our own implementation and interpretation of those theories to be reviewed, believed, changed, or rejected by as a result of continuing validation processes including peer assessments..

In this paper we report on one such "process validation" exercise–an effort at theoretical refinement focused on protest. We provide an account of tests of the implications of rival operationalizations that incorporate key elements of a) Timur Kuran's theory of political cascades and b) Peter Eisinger's theory of the structure of political opportunities. Specifically, in these experiments we ask about the effect of adjusting the routines constraining agent behavior associated by giving an important role to the mechanism of cascades of mobilization arising from changing perceptions of the behavior of others as that behavior impacts on an internal calculus to "protest" or not.

Before introducing and describing these experiments, and then discussing their results, it will be useful to map the kinds of variables listed above, associated in principle with any interest in political mobilization, to their current implementation in the model. Table 1 lists the six categories of variables of interest in the left hand column. The right hand column lists the corresponding locations in the model where that category of variable is either operationalized or subsumed within a stochastic distribution. Within this table the specific targets of the experimentation reported in this paper appear in boldface.

Theoretically Significant Variables	**Model Operationalizations**
Strain	Anger: difference between the return on investment of an agent's activated affiliation and the overall return on activation publicly associated with an available alternative affiliation
Organizational Frameworks for Mobilization	Groups of agents demarcated by shared activated affiliations and shared identities in agent repertoires; Categories within the Dynamic Political Hierarchy registering the status of groups in relation to the dominant group or coalition in the political system
Contexts conducive to bottom-up mobilization	Subsumed stochastically within the "Mobilization Factor"
Norms and Heuristics to Act	Rules operating on agents in groups categorized at any particular time within one or another level of the Dynamic Political Hierarchy; **(plus, in the "treatment" condition, a routine implementing agent consideration of number of visibly mobilized agents)**
Overcoming personal temptations to shirk associated with collective action theory	Subsumed stochastically within the "Mobilization Factor"
Vulnerability of regime	Endogenously produced configurations of the Dynamic Political Hierarchy, featuring stronger or weaker dominant groups and coalitions, affect amounts and types of mobilization by non-dominant groups **(plus, in a "treatment" condition, an increase in the likelihood of protest depending on the ratio of agent activated group size and the size of the dominant group)**

Table 1: Model Operationalizations of Theoretically Significant Variables

3 Testing Theories of Protest

The operationalizations noted in Table 1 are the result of multi-year processes of peer review, self-assessment, and self-re-assessment of model performance in relationship to expectations and to a "Ground Truth Data Set" compiled by Lockheed Martin, ATL, in connection with its work on the Defense Department's ICEWS program and as refined and checked by independent researchers. Our models, and specific elements within them, thus have their own dynamic life cycles. An important goal of this paper is to learn about and help discipline that process by focusing in a detailed way on the question of whether or not to adjust operationalizations by drawing on theory not previously encoded into our model. As noted, and as highlighted in Table 1, we will focus on a simplifying assumption we make in our operationalizations of protest. We test two well-established theories to see if they improve model performance by validating our results against empirical evidence.

This experiment can be located in the context of a broader discourse relating to agent-based simulation of collective mobilization. By considering Joshua Epstein's model of civil violence (Epstein, 2006). Epstein's model identifies four major components that affect an individual's likelihood to protest: hardship, legitimacy of the regime, the agent's risk aversion, and the probability of arrest. Of these components, his model endogenously and atheoretically sets the parameter values for hardship, legitimacy, and risk aversion. Probability of arrest is theoretically grounded in Kuran's cost/benefit approach to social movements in much the same way as it is in one of our experimental conditions. If hardship and legitimacy are understood as related to the overall categories of "strain" and "norms and heuristics," then it is apparent that a good deal of the machinery in Epstein's model that is exogenous to it is *endogenous* to ours. We are, however, here exploring the potential of adding the element in Epstein's model that is endogenous to our somewhat more elaborate existing model of political mobilization as it is embedded within specific country virtualizations. The current protest operationalization used in the models described above includes two necessary conditions in order for an individual agent to protest–political alienation (Organizational Frameworks for Mobilization from Table 1) and anger (Strain from Table 1)–as well as a random element called the "Mobilization Factor" that determines the likelihood of an agent actually protesting when the two necessary conditions are met. For the purpose of this study, we will hold constant the political alienation and anger requirements while we test two theories that relate to the Mobilization Factor.

Political alienation, the operationalization of which will be held constant across all experiments, is determined by a theoretical module embedded within our country models known as the Dynamic Political Hierarchy (DPH). A thorough review of the theoretical foundations and specific operationalizations of the DPH is available elsewhere (Lustick et al., 2012), but the basic premise of the module is that it computes all the overlapping group affiliations of the agents in a given country and uses those relationships to build a hierarchical network of relationships between all the groups in a specific country model at each moment in time. The DPH also calculates the politically dominant group at each moment of time. By combining information about the politically dominant group with the hierarchical network information, we are able to identify groups that are political allies of the dominant group, in opposition to the dominant group, or radically opposed to the dominant group. The agents that meet our moderate political alienation requirement, and are therefore able to protest, are members of groups in political opposition to the dominant group, but not radical opposition.

Anger, the operationalization of which will be held constant across all experiments, within our country models is based on a simple interpretation of relative deprivation theory.(Gurr, 1970) Stochastic perturbations to the model, known as biases, serve as random fluctuations in favorability scores for each of the groups in the model. At each time step all groups are assigned a score from within a constrained range, generally -3 to 3, which is intended to represent the environmental conditions that can either be positively or negatively affecting the group but which are not otherwise represented in the simulation. For an individual agent with an activated identity and a repertoire of latent identities, these bias scores can serve as a measurement of relative deprivation. In a perfectly responsive system, an agent would be able to immediately activate on the identity in its repertoire with the highest bias, the most favorable environmental conditions at the time. However, real-world systems are not, and models intended to simulate the real world should not be, perfectly responsive. Rather, surrounding influence structures and the "stickiness" of current identity affiliation tend to prevent agents from optimizing their activated identity based on changing bias scores. As a result, an agent can be activated on an identity with a low bias while having identities in its repertoire with higher biases that for whatever reason it is unable to activate on. When this is true, we consider the agent to be angry because of the disparity between its present state and its other possible states. This anger is the second necessary condition for an agent to protest.

The Mobilization Factor represents the likelihood that an agent will protest when the previously described conditions, political alienation and personal anger, are met. This stochastic factor is intended to capture theoretically significant variables not explicitly operationalized within our model (e.g. Contexts conducive to bottom-up mobilization and Overcoming temptations to shirk from Table 1). The Mobilization Factor can range in value anywhere from 0% meaning that agents will never protest, even when they are alienated and angry, to 100% meaning that agents will protest every time they are alienated and angry. Using a simple theoretical parameterization of Mobilization Factor we can hold the value constant. In the case of our baseline experiment it is held at 10%. In other words, we assume that individuals who are candidates for protest due to their alienation and anger will only actually carry out protests 10% of the time due to a variety of structural, psychological, political, or other conditions that impede or discourage protest. In this simple parameterization we do not seek to explain exactly what these conditions are or to measure them in the model, we simply claim that all of the various factors resolve to an average protest likelihood of 10%.

However, there is a large body of political science literature that attempts to explain some of the conditions that determine whether or not an individual that is inclined to protest actually acts to implement that inclination. It is the purpose of this study to compare two schools of thought within the protest literature by implementing the Mobilization Factor differentially according to specified tenets of each theory and comparing the results to empirical evidence from the real world.

4 Protest movements as political bandwagoning

As is widely understood, collective action theory, despite its elegance and the valuable insights it promotes, predicts much less collective action than we actually see in the world. In general, it seems easier, in many settings, including those involving large numbers of individuals, to induce more risk-taking or cost-incurring behavior (notably voting) than would be expected out of a pure expected utility calculation. Among the mechanisms that have been advanced to explain this discrepancy is Timur Kuran's theory of heterogeneous thresholds of preference falsification (Kuran, 1991). Kuran models a population as comprised of discontented individuals with very different propensities for risk-acceptance. Individuals decide whether their thresholds for accepting risks associated with public political mobilization are met by monitoring the actions of others. Using these ingredients, Kuran presents an intriguing and influential argument for why cascades toward large-scale protests can erupt rather suddenly, but rarely do, despite generalized discontent and individual fears of non-negligible retribution associated with an individual contribution to mobilization success which is negligible. Kuran proposes that cascading patterns of transformation — from collective apathy

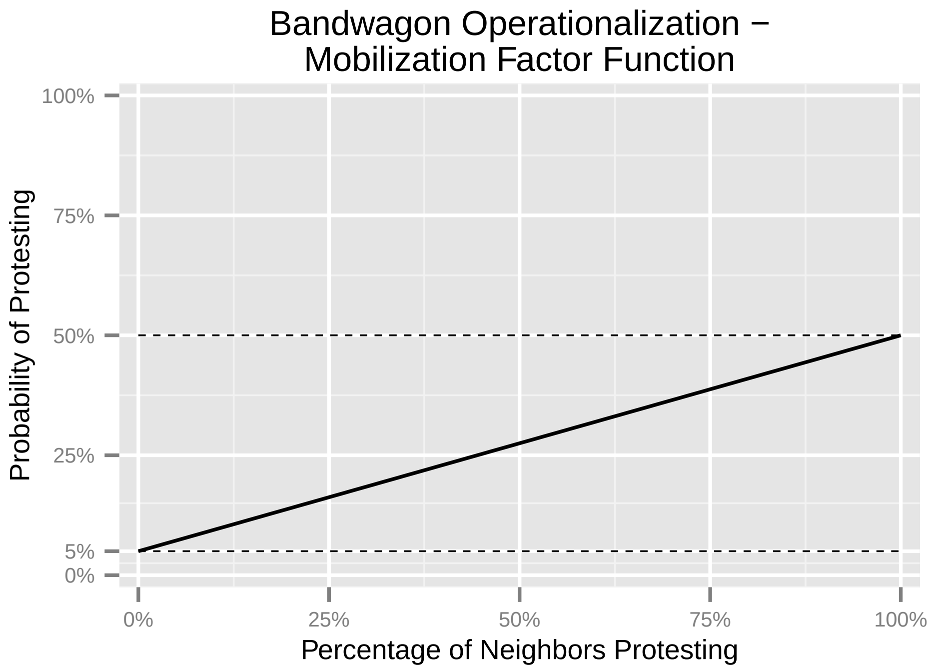

Figure 1: Mobilization Factor function for **Bandwagon operationalization**

toward acquiescence in, or support for, an authority structure, to collective defiance and explicit renunciation of its claims — are traceable to the catalytic effect of actions by individuals with low thresholds for tolerating psychic dissonance when many other individuals are near their thresholds

4.1 Operationalization

The basic premise of protest movements as political cascades is that an individual is more likely to protest if he or she observe others protesting. We have implemented this positive feedback mechanism in the model by increasing the likelihood of an agent protesting–the Mobilization Factor (or $P(Mobilizing)$ below)–as a function of how many other agents are protesting in the eligible agent's neighborhood. The specific function we have used to determine is mobilization factor is:

$$\begin{aligned} P(Mobilizing) = ((&MaximumMobilizationFactor - MinimumMobilizationFactor) \\ &*PercentageOfNeighborsProtesting) + MinimumMobilizationFactor, \end{aligned} \tag{1}$$

where $MaximumMobilizationFactor$ and $MinimumMobilizationFactor$ are the limits to the Mobilization Factor value and $PercentageOfNeighborsProtesting$ is the percentage of agents within a distance of four that are protesting. $P(Mobilizing)$ is the probability that the agent will mobilize as a protest agent in the current timestep. In our experiment, we set $MinimumMobilizationFactor$ to 0.05 and $MaximumMobilizationFactor$ to 0.5. Figure 1 shows the relationship between these key variables.

5 Protest movements as opportunistic or structural phenomena

Peter Eisinger, one of the early proponents of this school of thought, considered the "structure of political opportunities" as the key element in determining when protest does or does not occur (Eisinger, 1973). He defined political opportunity structure as "the degree to which groups are likely to be able to gain access to power and to manipulate the political system." More recently, McCarthy, Britt, and Wolfson observe that "the elements of the environment have manifold direct and indirect consequences for people's common decisions about how to define their social change goals and how to organize and proceed in pursuing those goals" (McCarthy et al., 1991). In other words, the political opportunity structure at the macro level, rather than individual or group agency is the major explanatory variable for social movements. From this angle we would expect to see protest movements more often in cases where the aggrieved population has reason to believe the movement has a significant chance of being successful and less often when the movement is likely to be quickly contained. Authors have applied concepts from this school of thought to explain a range of non-violent and violent movements including the American women's movement (Costain, 1992), to peasant mobilization in Central America (Brockett, 1991), and insurgency (Fearon and Laitin, 2003).

5.1 Operationalization

In order to implement a simple understanding of the structure of political opportunities affecting the likelihood of protest movements within our models, we will use concepts from the Dynamic Political Hierarchy (DPH) described earlier in this paper. At each time step in the model the DPH determines which group is politically dominant, which groups are political allies, and which groups are in the opposition or radical opposition. These last two groups, the opposition and the radical opposition, are the only groups eligible to protest in the country models used for this experiment. To operationalize this theory we have started with a basic assumption that the prospects of "success" for a protest movement depend on the ratio of influence between the summed influence of (a) the activated identity of the agent considering whether or not to protest and (b) the dominant political group. When the dominant group is large and opposition group is small, the likelihood of a successful protest movement is small and therefore the likelihood of any individual agent protesting, the Mobilization Factor, is small. As the dominant group shrinks or the opposition group grows, the Mobilization Factor increases. The specific function we used to determine the mobilization factor is:

$$\begin{aligned} P(Mobilizing) = ((&MaximumMobilizationFactor - MinimumMobilizationFactor) \\ &*RatioOfActivatedIdentitytoDominant) + MinimumMobilizationFactor, \end{aligned} \tag{2}$$

where $MaximumMobilizationFactor$ and $MinimumMobilizationFactor$ are the limits to the Mobilization Factor value and $RatioOfActivatedIdentitytoDominant$ is the ratio between the summed influence of a given agent's activated identity and the summed influence of the dominant identity. $P(Mobilizing)$ is the probability that the agent will mobilize as a protest agent in the current timestep. In this experimental condition (as in the other condition above), we set $MinimumMobilizationFactor$ to 0.05 and $MaximumMobilizationFactor$ to 0.5. Figure 2 shows the relationship between these key variables.

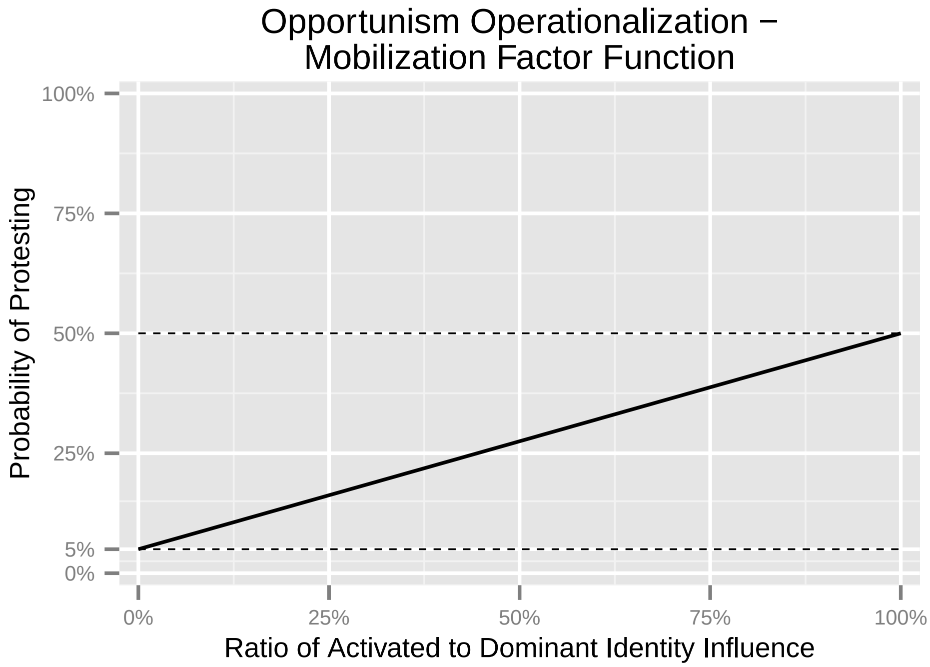

Figure 2: Mobilization Factor function for **Opportunism operationalization**

6 Experimental Design

Our experiment was designed to test our operationalizations of the two theories discussed above using six different country models we have developed as part of the W-ICEWS and V-SAFT projects funded by the Human Social Cultural Behavioral (HSCB) program of the Department of Defense, administered by the Office of Naval Research. The target countries were Bangladesh, Egypt, Pakistan, Philippines, Thailand, and Venezuela. For each model we produced 100 unique, simulated trajectories of 52 time steps (1 time step is calculated to represent one week) for three conditions: 1) A baseline where $P(Mobilizing)$ is kept constant at 10%, 2) An operationalization of the political cascade (bandwagoning) theory where $P(Mobilizing)$ ranges from 5% to 50% based on the number of protests in an eligible agent's neighborhood ($PercentageOfNeighborsProtesting$), and 3) An operationlization of the structural explanation of protest where $P(Mobilizing)$ ranges from 5% to 50% based on the ratio of influence of the dominant political group to the influence of opposition groups ($RatioOfActivatedIdentitytoDominant$).

We will compare the results from the three experimental conditions across six country models with 12 years of event data gathered under the W-ICEWS program. Specifically, we will compare the distribution of our protest results to "medium hostility" events in the event data. The events are coded using an event coding method provided by the ITRACE system within W-ICEWS developed by Raytheon BBN Technologies. Each event is coded to a particular CAMEO code (Gerner et al., 2002) which in turn corresponds to a particular Goldstein score (Goldstein, 1992). The Goldstein score is meant to measure the intensity of a given event on the conflict-cooperation dimension. Each event is also coded to a particular country when both the source and target actors are located within that country's dictionary. All events that have a Goldstein score between -8 and -4 are then coded as “medium hostility” and summed for a particular country/day in the dataset. Examples of CAMEO codes more medium hostility events include “Protest”, “Threaten”, or “Demand’ and include sub-types such as “Engage in violent protest” or “Mobilize or increase police power”. For more information on the ITRACE system, see Lautenschlager et al. 2012.

The purpose of this comparison was to see, for our admittedly limited sample of six countries, whether a shift from our baseline assumption of holding the probability of mobilizing at 10% to one of our theoretical operationalizations tend to produce distributions of protest events that better reflect the available empirical data.

7 Results

Below is a description and discussion of the results from our experiment as well as empirical data from the ITRACE program.

7.1 Describing the Data

Figure 3 on page 10 shows the actual weekly totals of medium hostility events in our six target countries. We can see that each country has a unique medium hostility trajectory over the past twelve years. We can also see that the y-axis varies dramatically from country to country. On the other hand, our experimental results are much more difficult to visualize. Given the significance of chaos effects, and in general, the impact of factors operating below the analytic horizon, no model of the future of a complex political system can identify the *actual* trajectory it will follow, only a distribution of possible, plausible, and probable futures. Accordingly, instead of one line, we have 100 lines per country-experiment. Some runs have very low medium hostility values for the entire year while others have dramatic spikes and plateaus. The values also vary widely from very low in some countries and runs to very high in other countries or particular episodes during the simulation.

In order to better understand what the shape of each distribution is like, we can view them in scatter plots by comparing each medium hostility value to the value from the prior week. This gives us a sense of the magnitude and direction of changes in medium hostility from week to week. Figure 4 on page 11 shows these plots for our six target countries as well as a line representing “y equals x” to show where the values would be changing to the same value from week to week. There are several interesting indicators we could extract from this type of visual, including the autocorrelation and delta. The lag-one autocorrelation is the simple Pearson’s correlation of the x and y axis of these charts.

We can see from the visuals that Pakistan’s values are clustered more closely around the ”y equals x” line, indicating a higher correlation between weekly medium hostility event totals. Venezuela, on the other hand, has more points lying further away from ”y equals x” line, indicating lower week-to-week correlation and a higher frequency of sudden shifts from high to low (or vice versa) in medium hostility values in Venezuela than in Pakistan. These data will be analyzed more deeply below by focusing on the absolute value of the distance from the diagonal, which would represent the magnitude and direction of change from week to week. If there are many points far from the line, we would expect a high average delta, meaning that change is not predictable.

7.2 Preparing the Data

In order to more clearly analyze and describe our results, we have converted all medium hostility values in the empircal data and experimental output into percentages of the maximum values for each particular dataset. This means that the maximum value for each country or experiment is 1.0

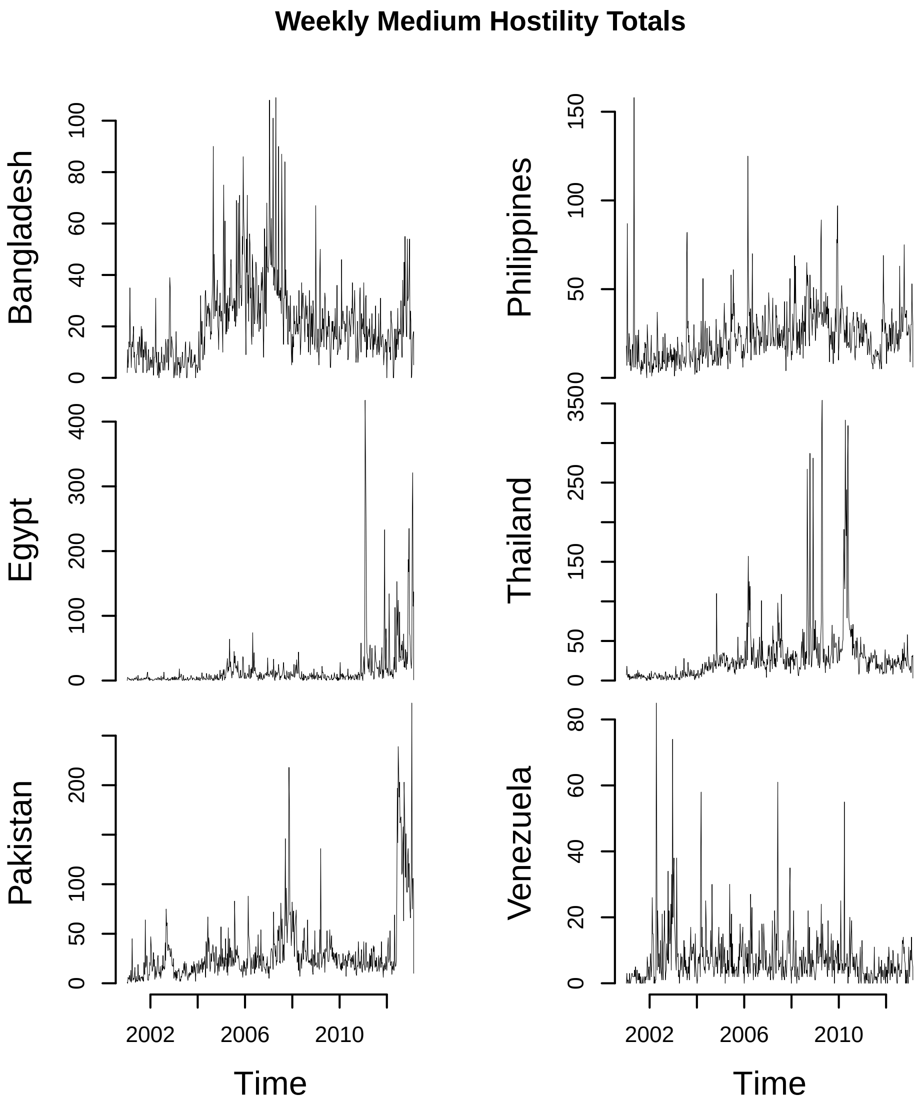

Figure 3: Weekly medium hostility totals in target countries

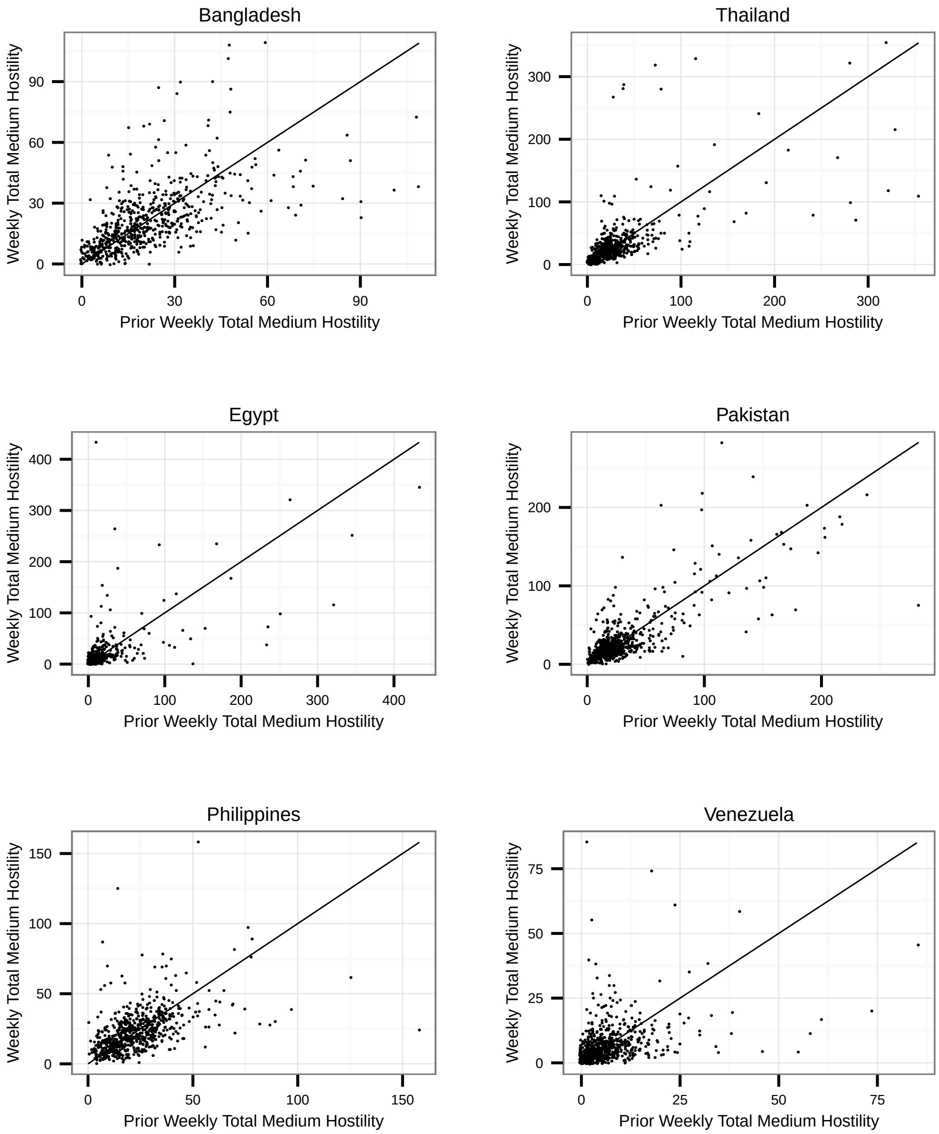

Figure 4: **Empirical data** - Scatter plots of the prior week (lag one) protest compared to the current week of protest during the twelve year period. (A slight "jitter" is applied so points overlap less often.) The line represents the "y equals x" line, indicating where the values would land if each week's value was equal to the last.

or 100%. A value of 0.5 or 50% would represent a value of half of the maximum. This conversion allows us to compare the relative shape of different distributions even when they have very different actual values. For example, the average number of medium hostility events in a week in Venezuela was 7.3 between 2001 and 2013 whereas Bangladesh had an average of 21.2. This scaling technique allows us to study the "shape" of the distribution in space and time by asking questions such as "how often does medium hostility increase from the middle of the distribution to the very high end of the distribution" instead of "how often does medium hostility increase from 10 to 20". [2]

7.3 Experimental Results

7.3.1 An example: Bangladesh

In Figure 5, we compare the Bangladesh empirical data with the three experimental cases outlined in section 6. We have labeled the first operationalization "Opportunism", described in section 5 and the second "Bandwagon", described in section 4. One thing we notice right away is that the experimental charts have many more points, because we ran 100 "years" of for each experiment, whereas we only have twelve years of data from the target country. We can also see that there are different shapes to each plot, with more points along the bottom and left axis in the Opportunism chart specifically, indicating a larger delta (distance from the diagonal). This is something that we do not see in the empircal data. This would indicate that the Opportunism operationalization tends to over-emphasize the likelihood of large negative and positive differences from the previous week's medium hostility values. On the other hand, we see that there are more values near the diagonal in the Opportunism case, whereas the Bandwagon experiment exhibits more points at a medium distance.

7.3.2 Comparing the Empirical and Experimental Data

In Figure 6 (page 14) and 7 (page 15), we can see the comparison of means for each country and experimental condition to the empirical data for the lag-one autocorrelation and one-week delta. In both figures, we are comparing the empirical value for each country to the values generated from each of the three experimental conditions. We then plot the absolute value of the difference between the empirical and experimental values to show how similar or dissimilar the experiment was to the real world in terms of weekly deltas and lag-one autocorrelation. One thing we can notice right away is that Pakistan seems to be an outlier in both cases, with experimental indicators very far from the empirical indicators. Also, neither operationalization clearly outperforms the baseline experimental

[2] In our model output, certain models produced wildly different actual values of medium hostility compared to the empirical data, which made it difficult to compare results. There could be a number of reasons for this discrepancy.

1. Certain models are weaker or stronger than others, and produce values that do not correspond well to the target country.
2. The empirical data represents only one possible history for a given country, and given different circumstances, a much different past could have occurred.
3. The baseline models we have chosen to use are tuned to the recent past, and may not be able to taken into account structural shifts in a country's politics (coups, elections, etc.).
4. Particular parameters used for a given operationalization (e.g. Bandwagon or Opportunism) produced medium hostility values that were much higher or lower than the baseline, and different values would have produced results closer to the distribution of the baseline experiment.
5. Our experiment only runs the country for one year and increasing the model run time would change the distributional nature of the results.

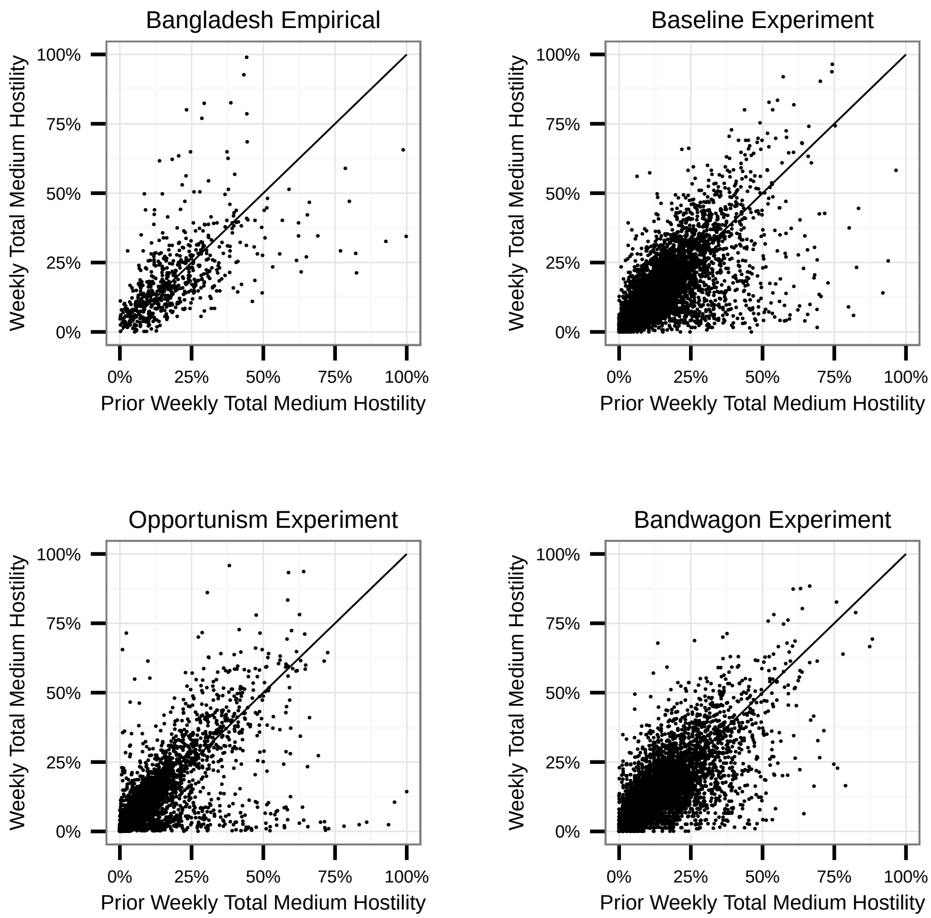

Figure 5: Scatter plots of the prior week (lag one) protest compared to the current week of protest in the empirical and three experimental conditions. The line represents the "y equals x" line, indicating where the values would land if each week's value was equal to the last.

condition. We can see that although Opportunism does underperform on most categories, it does perform best in the Egypt delta case particularly, and often either outperforms the Bandwagon case compared to the Baseline or is relatively close. The Bandwagon and Baseline cases often showed mixed results or are very close. The order of the three cases is also not always the same between the two figures, indicating that certain operationalization tends to capture only a key aspect of the distribution.

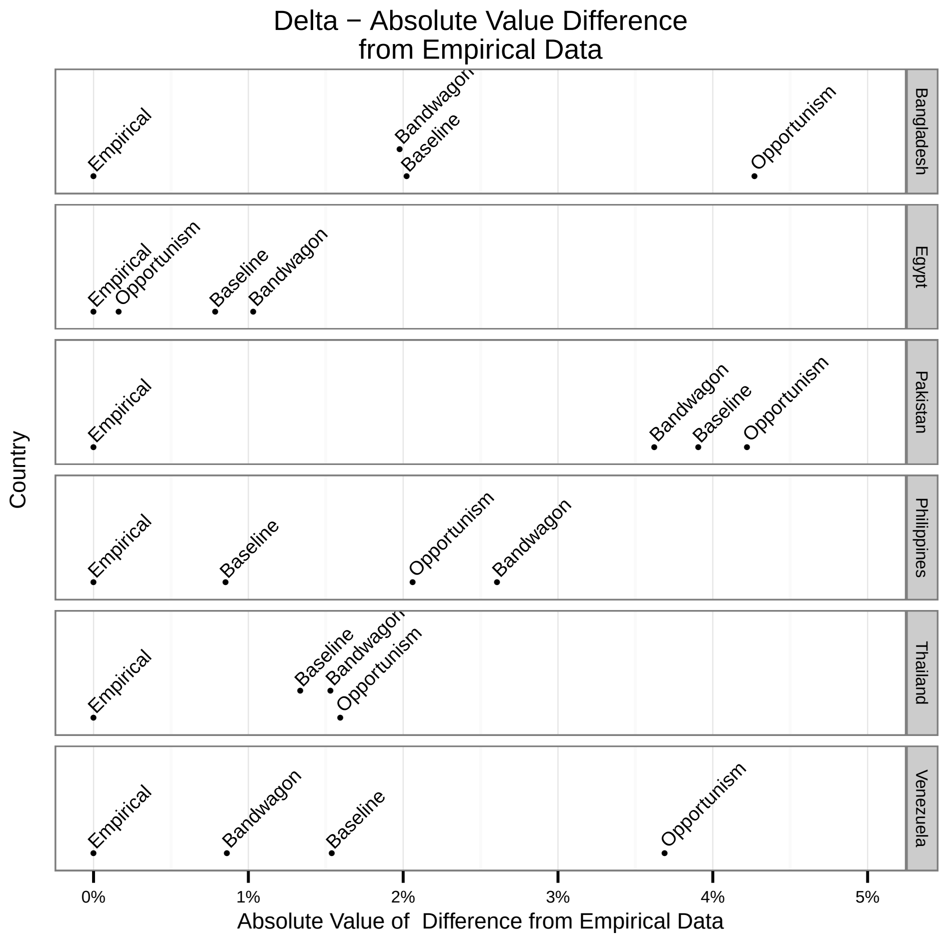

Figure 6: Absolute value of the difference of means of one-week delta from empirical and experimental data.

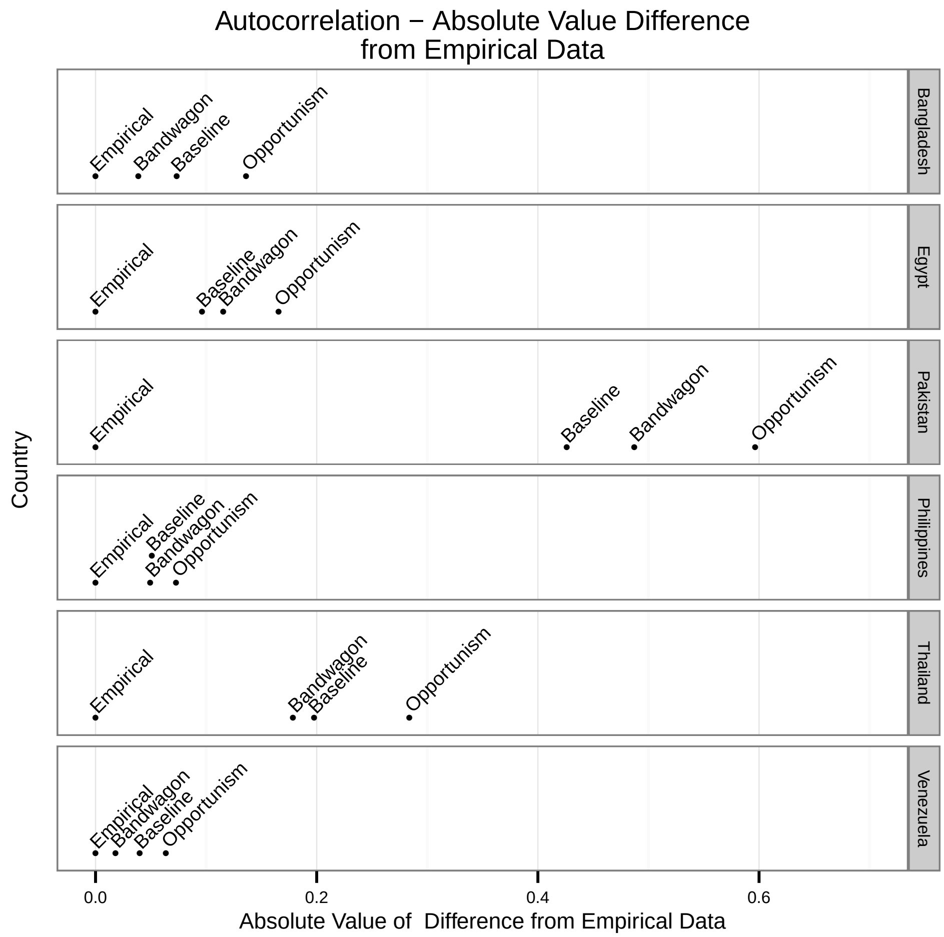

Figure 7: Absolute value of the difference of means of lag-one autocorrelation from empirical and experimental data.

7.3.3 Aggregation of All Data

Lastly, we might wonder what the plot might look like if we aggregated all countries together into one larger scatter plot, which we can see in Figure 8. Overall, it is worth noting the general morphological similarity exhibited by all four of these plots, namely a dense mass of values clustered around the diagonal in the bottom left quadrant, accompnied by attenuating distributions of values into the other quadrants. This suggests a certain degree of face validity to our models, even taking into account the clearly flawed Pakistan model, since the contours of the state space adumbrated by the limited empirical data available are replicated in our models by each condition. We also note a relatively high number of values in our experimental conditions lying close to the x-axis, which represent drastic outcomes from high levels to very low levels of mobilization. This pattern can be interpreted as an artifact of the non-linearities in the models, or as illuminating the greater than empirically expected possibility of these multi-sigma events.

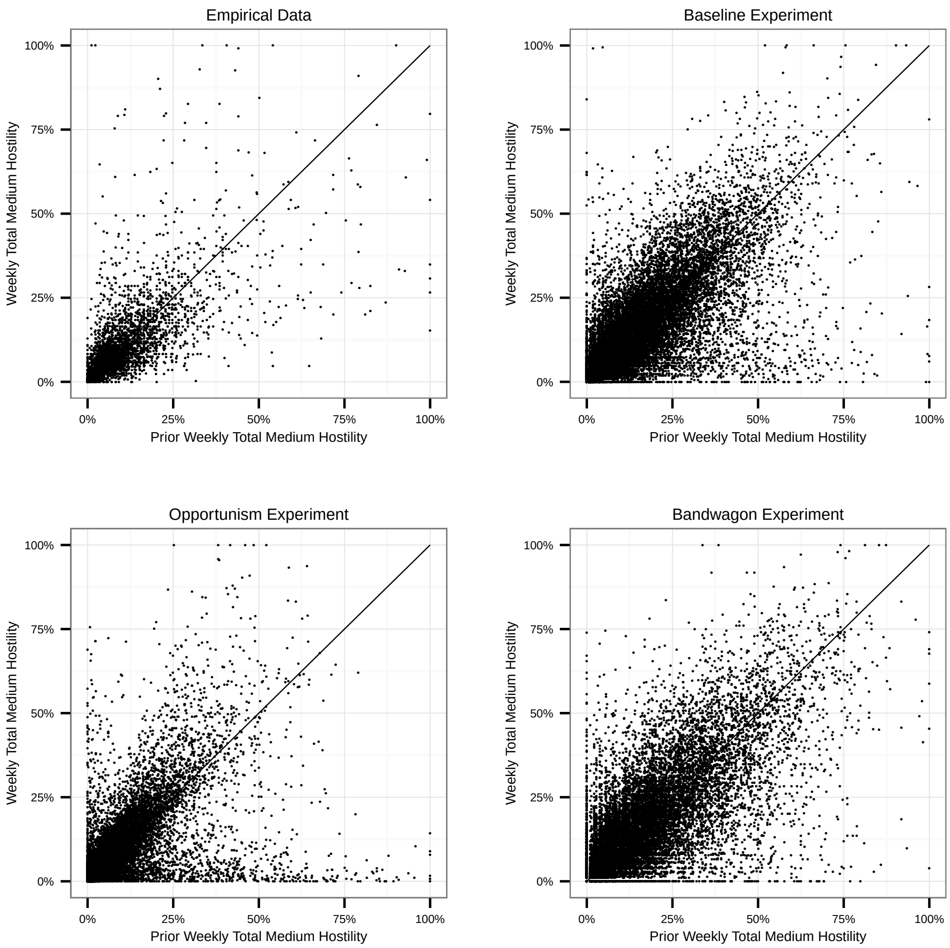

Figure 8: Comparison of prior week to current week values in all countries in empirical and experimental data.

8 Conclusion

The purpose of this paper has been to detail the very challenging task of empirically validating computational social science models. Admittedly, our aim of empirical validation may have been overambitious and our results are inconclusive, but we hope to contribute by identifying major challenges that face the computational modeling community and sharing our preliminary methods for overcoming some of these challenges. The key challenges as we see them include:

1. **Comparing distributional model results to a single empirically true history of the real world**

   One of the major strengths of stochastic agent-based models like the ones we used in these experiments is the fact that they produce a distribution of possible outcomes given a certain set of model specifications. This allows us to analyze patterns of what is possible, plausible, and probable for a certain model and provides a more accurate projection of the uncertain future of the real-world than a model that provides a single point prediction, but it is also much more difficult to validate empirically. One of our 100 model trajectories could very closely match empirical findings in the real world or 90 out of our 100 model trajectories could closely match empirical findings, but neither case definitively informs us about the accuracy of the model. Everything depends on how likely that real world trajectory was in comparison to the distribution of once possible but unrealized trajectories.

   In this experiment, we attempt to deal with this problem by effectively increasing the number of real-world events to which we can compare our model results. Instead of looking at one 12-year time series of medium hostility events, we look at the 633 week-to-week changes featured within that trajectory and compare that to the 5100 week-to-week changes in our simulated distribution. This allows us much more flexibility in terms of the metrics we use to compare the datasets, though autocorrelation and delta values of week-to-week change still proved difficult to draw definitive conclusions from in this case.

2. **Trade-offs between model complexity and empirical relevance**

   Decisions about the appropriate levels of complexity feature major trade-offs in moving either way on a continuum from very simple models to very complex models. The advantage of using a simple model is that it is much easier to isolate particular parameters and test for the effect. We could have built a model that featured little more than an operationalization of political bandwagoning or structural opportunism, and ignored other elements like political alienation, anger, or complex identity repertoires. This type of model would have produced much clearer results, but would be hard to validate because so many elements of the real-world need to be ignored in order to build a simple model. On the other hand, the complex models we chose to use have a stronger relation to the real world and therefore some opportunities for empirical validation, but also are difficult to experiment with in a clear, precise way and produce results that are more challenging to trace to individual, isolated variables.

   There is no universal answer about the appropriate level of model complexity. Instead, it is determined by the goals and needs of the modeler. Our goal is to simulate political phenomena in particular countries as best we can, so we have opted to use more complex models that we hope to satisfactorily compare to empirical findings. However, this empirical validation is a constant work in progress and we hope this paper is a first step that we can build upon as we learn from our errors and develop new techniques.

3. **Establishing a sufficient sample size**

> Building virtualized agent-based models of countries is a time-consuming task. In this experiment, we use a sample of only six countries with twelve years of empirical data. While we believe this to be a sufficient base from which we can start experimenting, refining techniques, and drawing conclusions, it is hardly a large enough sample to draw definitive conclusions about the merit of broad theories of protest or the validity of our modeling techniques. We hope to double the number of countries we are modeling within the next year and the ITRACE data we are using for empirical validation is being updated on a monthly basis, so our sample size will grow. In the meantime, it is important to be cautious about the conclusions we draw and recognize that we are still only exploring a small portion of the world.

In addition to the challenges we've identified, it is important to point out the four preliminary findings from this experiment that will affect our work going forward. First, neither the bandwagoning nor the political opportunism operationalizations of protest convincingly outperformed our baseline model. Based on this conclusion and our general operating principle of adding complexity to the model only when it improves model performance or theoretical clarity, we will not be adding either operationalization to our standard country models. Second, this method of comparing competing theories by operationalizing them within a model and comparing different output distributions to empirical data can provide significant insights into the strengths or weaknesses of the theories. For this to be the case one must a) have confidence in the underlying model, b) test a sufficient range of operationalizations for each theory that do indeed reflect the central tenants of the theories, c) apply the operationalizations to a range of cases large enough and diverse enough to avoid artifactual findings, and d) have high confidence in the empirical data. Third, the exercise of implementing distinctive operationalizations for bandwagoning and opportunism has led us to consider closely the elements of these theories that are present in our basline condition, thereby improving the theoretical scaffolding of our modeling effort. Fourth, under all three conditions, Pakistan was the worst performing model in relation to its empirical metrics. We will be revisiting the Pakistan model in order to identify areas for improvement.

References

Brandon Alcorn, Miguel Garces, and Allen Hicken. VirThai: A PS-I implemented agent-based model of Thailand as a predictive and analytic tool. *Stanford Journal of East Asian Affairs*, 12 (1), 2012.

Charles D Brockett. The structure of political opportunities and peasant mobilization in central america. *Comparative politics*, 23(3):253–274, 1991.

Kanchan Chandra. *Constructivist Theories of Ethnic Politics*. OUP USA, 2012.

Paul Collier and Anke Hoeffler. Greed and grievance in civil war. *Oxford economic papers*, 56(4): 563–595, 2004.

Anne N Costain. *Inviting Women's Rebellion: A Political Process Interpretation of the Women's Movement*. Johns Hopkins Univ Pr, 1992.

Peter K Eisinger. The conditions of protest behavior in American cities. *The American Political Science Review*, pages 11–28, 1973.

Joshua M. Epstein. *Generative Social Science: Studies in Agent-Based Computational Modeling*. Princeton University Press, 2006.

James D Fearon and David D Laitin. Ethnicity, insurgency, and civil war. *American political science review*, 97(1):75–90, 2003.

Deborah J Gerner, Philip A Schrodt, Rajaa Abu-Jabr, and Ömür Yilmaz. Conflict and mediation event observations (CAMEO): A new event data framework for the analysis of foreign policy interactions. *International Studies Association, New Orleans*, 2002.

Joshua S Goldstein. A conflict-cooperation scale for weis events data. *Journal of Conflict Resolution*, 36(2):369–385, 1992.

Ted R Gurr. *Why Men Rebel*. Princeton University Press, 1970.

Timur Kuran. Now out of never: The element of surprise in the east european revolution of 1989. *World politics*, 44(01):7–48, 1991.

Jennifer Lautenschlager, Rachel Hingst, James Starz, Mark Hoffman, Joe Roberts, Jason Losco, and Kevin Spivey. Supporting situation understanding (past, present, and implications on the future) in the USSTRATCOM ISPAN Program of Record. *Advances in Design for Cross Cultural Activities*, page 181, 2012.

Ian S Lustick, Brandon Alcorn, Miguel Garces, and Alicia Ruvinsky. From theory to simulation: the dynamic political hierarchy in country virtualisation models. *Journal of Experimental & Theoretical Artificial Intelligence*, 24(3):279–299, 2012.

John D McCarthy, David W Britt, and Mark Wolfson. The institutional channeling of social movements by the state in the united states. *Research in Social Movements, Conflicts and Change*, 13(2), 1991.

Neil J Smelser. *Theory of Collective Behavior*. Routledge, 1962.