IDA

INSTITUTE FOR DEFENSE ANALYSES

Force Sizing for Stability Operations

Main Report

R. Royce Kneece Jr., Project Leader
David A. Adesnik
Jason A. Dechant
Michael F. Fitzsimmons
Arthur Fries
Mark E. Tillman

March 2010

IDA Paper P-4556

Log: H 10-000408

The Institute for Defense Analyses is a non-profit corporation that operates three federally funded research and development centers to provide objective analyses of national security issues, particularly those requiring scientific and technical expertise, and conduct related research on other national challenges.

About this Publication

This work was conducted by the Institute for Defense Analyses (IDA) under contract DASW01-04-C-0003, Task BA-6-3005, "Sizing Force Requirements for Stability Operations," for the Office of the Director, Cost Assessment and Program Evaluation. The views, opinions, and findings should not be construed as representing the official position of either the Department of Defense or the sponsoring organization.

Acknowledgments

The authors wish to thank the reviewers, Dr. David Graham, Mr. Michael Leonard, Mr. John Shea, Mr. Vincent Roske, and Mr. Michael Dominguez for valuable comments and suggestions. We also obtained helpful guidance and suggestions from our sponsor representatives, Dr. John Crino and Mr. Andrew Caldwell. We are also indebted to a number of people in IDA with whom we had valuable conversations, in particular: Dr. John Cann, Dr. Colin Doyle, and Mr. Richard Polin. We also wish to thank numerous people and organizations outside IDA for valuable discussions and making data available to us, especially: Center for Army Analysis—Mr. Touggy Orgeron, Ms. Justice Blaho, and LTC David Smith; Defence Science and Technology Laboratory, United Kingdom Ministry of Defense—Dr. Andrew Hossack; U.S. Army—Brigadier General Herbert R. McMaster, Jr., Lieutenant Colonel (ret.) Douglas Ollivant, Colonel (ret.) Conrad Crane, and Colonel (ret.) Peter Mansoor; U.S. Marine Corps – Colonel Julian "Dale" Alford.

INSTITUTE FOR DEFENSE ANALYSES

IDA Paper P-4556

Force Sizing for Stability Operations

Main Report

R. Royce Kneece Jr., Project Leader
David A. Adesnik
Jason A. Dechant
Michael F. Fitzsimmons
Arthur Fries
Mark E. Tillman

EXECUTIVE SUMMARY

This study was conducted for the Office of the Director, Cost Assessment and Program Evaluation (CAPE) to develop methodologies for estimating force requirements for large-scale stability operations. The goal of the study was to provide Department of Defense (DoD) leadership with a rough idea of the capabilities and limitations of the programmed ground force structure to conduct stability operations in a broad range of countries. The estimates from this study are intended for use in force planning and evaluation within the context of the DoD Future Years Defense Program, not *operational* planning.

The term "stability operation" encompasses a wide range of military activities. However, operations that involve large-scale counterinsurgency (COIN) operations are, historically, the most demanding for programmatic force sizing purposes and were the focus of the study. Background for the current study is found in work IDA performed in late 2005 supporting the 2006 Quadrennial Defense Review (QDR). Although in 2005 there was a paucity of data on historical stability operations, more recently, considerable additional research and data collection have been conducted. Specifically, the study obtained a database recently compiled by the Center for Army Analysis and extracted 41 conflict cases involving COIN operations for its analysis.

In 2006 the U.S. Army and Marine Corps developed a joint doctrinal field manual, FM 3-24, *Counterinsurgency*, that provides important guidance on force sizing for COIN operations. The manual suggests figures for "force densities" (troops per 1,000 inhabitants in the area of operations) required for effective operations--for example, 20 troops per 1,000 is cited as a minimum requirement. That figure has become a widely-referenced rule of thumb. This study sheds light on the evidence supporting that guidance, confirming the 20 troops per 1,000 figure as a *minimum*. However, the field manual also implicitly suggests 25 troops per 1,000 as the upper end of a range—a figure *not* confirmed by this study. We found that force densities of 40-50 troops per 1,000 may be required for high confidence of success.

The study reinforced the findings cited above by employing statistical analyses (logistic regression) on the selected historical data, finding a statistically significant relationship between force density and conflict outcomes for COIN operations. Since

these findings are at odds with the findings of other research organizations, the reasons for the differences in results were also investigated by the study team. The causes were three-fold: (1) we computed force densities using the populations in the *actual area of military operations*, whereas most other studies used populations for the entire country; (2) we categorized an operation as a "success" if the counterinsurgency force was *not defeated* militarily (other researchers used broader criteria including political outcomes), and (3) we scored certain conflicts as "indecisive" (and thus a "success" militarily) that others scored as a "loss." Under these conditions, we found that the logistic regression provided a coefficient for the force density independent variable with a *p-value* of less than 2%; p-values of 5% or less indicate a statistically significant relationship. The resulting regression equation provides an estimated probability of success of 50% for a force density of 16 troops per 1,000, and a probability of success of 75% for a force density of 40 troops per 1,000.

Data from recent operations in Iraq and Afghanistan also provided additional corroboration with regard to the relationship between force density and campaign success. The peak of the surge in Iraq achieved a force density for the total counterinsurgency force (U.S., coalition forces, and Iraqi forces) of 20 per 1,000 based on the population of the entire country—higher (but undetermined) for the actual area of operations. In Afghanistan, on the other hand, force densities achieved during the period covered by the study were much lower.

Drawing on the analyses summarized above, the study developed several techniques for estimating future force requirements for COIN-like stability operations and applied them to postulated operations in 54 countries. Using different approaches, estimates were found to vary by factors of two or more. Thus, projecting force requirements for future stability operations is subject to a large degree of uncertainty. Using mid-range estimates, the current U.S. ground force posture could probably sustain a successful COIN-like stability operation in most of the countries considered, but several key countries with larger-populations would likely be infeasible, unless the area of operations within the country could be limited significantly.

Although beyond the scope of this study, it is important to note that developing the right strategy and tactics is at least as important as force size, and the political landscape is an overarching factor.

The Classified Appendix identifies the countries for which force size projections were made and discusses applications of the methodologies to stability operations associated with the Defense Planning Scenarios.

The study developed the following recommendations based on its analysis:

- **Revise Army Field Manual 3-24 to reflect more recent analyses of historical data since its publication in 2006.**
- **Conduct additional research of historical data.** Given the sensitivity of our results to factors discussed above, more research into historical data is warranted.
- **Capture, organize, and archive data from ongoing operations.** Although this study was able to draw on the vast amount of data collected from operations in Iraq and Afghanistan, we did not find the data to have been organized, categorized, and archived in a manner that facilitates analysis. Those data are an extraordinarily valuable resource for future analysis of stability/COIN operations, but the value will be significantly lessened if the data are not better organized and archived with appropriate controls over access.
- **Define demands for a "whole of government" approach to COIN-like stability operations.** This study, by design, focused solely on military force size; but meeting demands for a broader range of services, including security, public safety, economic, political, justice, social and public welfare, are also of primary importance. A follow-on effort would analyze the appropriate role of the military in the "whole of government" effort and the relationships between military and civilian activities.

CONTENTS

APPENDICES

FIGURES

1. OVERVIEW

A. INTRODUCTION

Stability operations are a broad set of tasks, missions, and activities conducted in concert with all elements of national power to maintain or reestablish security and basic services to areas beset by disaster and strife. Department of Defense Instruction (DoDI) 3000.05 establishes "stability operations [as] a core U.S. military mission that the Department of Defense [DoD] shall be prepared to conduct with proficiency equivalent to combat operations."[1]

DoDI 3000.05 also directs the Office of Director, Cost Assessment and Program Evaluation (CAPE), to "assess the sufficiency of resources related to irregular warfare and stability operations within the Future Years Defense Program."[2] The Institute for Defense Analyses (IDA) was asked to conduct this study to assist in that task. The purpose of the study is provide DoD leadership with a rough idea of the capabilities and limitations of the programmed ground force structure to conduct stability operations in a broad range of countries. The estimates of this study should not be used for *operational* planning (though they may serve as a starting point); rather their intended use is for programmatic force planning and evaluation.

The term "stability operation" encompasses a wide range of military activities; however, for programmatic force-sizing purposes, operations that involve large-scale counterinsurgency (COIN) or COIN-like operations historically are the most demanding for military forces, and thus will be the focus of this study.

Background for the current study is found in work IDA performed in late 2005. In supporting the 2006 Quadrennial Defense Review (QDR), the then Office of Director, Program Analysis and Evaluation (now CAPE) asked IDA to conduct a quick-turn analysis to produce broad-brush estimates of ground force size requirements for stability operations in selected countries of interest. The objective was to apply some "outside-the-box" thinking about potential force demands for a broad range of stability operations, in

---

1 Department of Defense Instruction (DoDI) 3000.05. "Stability Operations." September 16, 2009, p. 2.

2 Ibid. p. 7.

contrast, conceptually, to the much higher resolution analyses of force sufficiency being conducted within the Department's "Operational Availability" series of studies. (Those studies were limited to the few scenarios in the set of approved Defense Planning Scenarios.)

In light of the continuing importance of stability operations and in preparation for the 2010 QDR, CAPE asked IDA to update and refine the 2005 analysis. Specifically, IDA was tasked to revise and expand the force size estimating techniques developed in 2005 by reviewing the most current research on stability operations; by incorporating lessons learned from ongoing stability operations; and by applying statistical analyses, where possible. In response, IDA evaluated and updated the original methodology and developed alternative approaches based on statistical analysis of historical data. This paper presents the results of those investigations.

In 2005 there was a paucity of data on historical stability operations. Fortunately, in the ensuing years, considerable additional research and data collection have been conducted. In particular, this study was able to capitalize on a major undertaking by the Center for Army Analysis (CAA) to compile a database of over 100 "irregular warfare (IW)" conflicts since World War II. From that data set, our study team extracted 41 conflict cases—all those that met the two criteria of (1) being a large-scale COIN operation and (2) having sufficiently complete data on COIN force sizes.

In addition, in December 2006 the U.S. Army and Marine Corps published a joint doctrinal field manual (FM) that provides important guidance on force sizing for counterinsurgency operations (the most demanding of stability operations).[3] The field manual has had a profound impact on the study and practice of irregular warfare. With regard to force sizing, the manual has exerted great influence via its suggestion that effective operations require the deployment of 20 troops per 1,000 inhabitants (known as the "force density") in the area of operations. That figure has become a rule of thumb not just in the military community, but also among Government officials and journalists concerned with critical national security issues, such as the decision to deploy additional troops to Afghanistan. Our study casts light on the evidence supporting the 20 per 1,000 figure, which has received little attention so far.

COIN operations comprise several lines of effort, such as protecting the population, defeating hardcore insurgents, developing the economy, building political and

---

3 FM 3-24 / MCWP 3-33.5 Counterinsurgency, Headquarters Department of the Army, December 2006.

social institutions, etc. This study, by design, focused solely on force size. Although not within the scope of this study, meeting demands for a broader range of services, including security, public safety, economic, political, justice, social and public welfare, are also of primary importance.

B. FINDINGS

1. Evaluation of 2005 Analytical Methodology

The methodology used in the 2005 analysis to project force sizes for stability operations was based on a series of scaling factors, such as population, area, a stability index, etc. for the target country, and a baseline reference operation (such as Operation Iraqi Freedom). As required, this study updated the methodology. An advantage of the approach is that one can use a reference operation that is well-understood and believed relevant for future operations. On balance, however, we reached the conclusion that, while the methodology served a useful purpose at the time, it has significant drawbacks from an analytical perspective. In particular

- It depends on factors whose statistical relevance to successful outcomes in more demanding stability operations (e.g., COIN) cannot be established
- It depends on subjective factors that lack a sound analytical basis and that are difficult to formulate so as to be mutually independent

Although we developed and present results using the (updated) 2005 methodology, methods based on data that were not available in 2005-2006 were judged more sound analytically, and were thus selected for our baseline projections.

2. Analysis of Historical COIN Data

As part of the study, IDA reviewed several databases of historical conflicts to determine those most appropriate for analysis. The best source was the CAA IW database. From that database, we selected a set of 41 historical COIN conflicts for analysis with the objective of establishing a relationship between force density and outcomes. We categorized each outcome as a success, a failure, or indecisive.

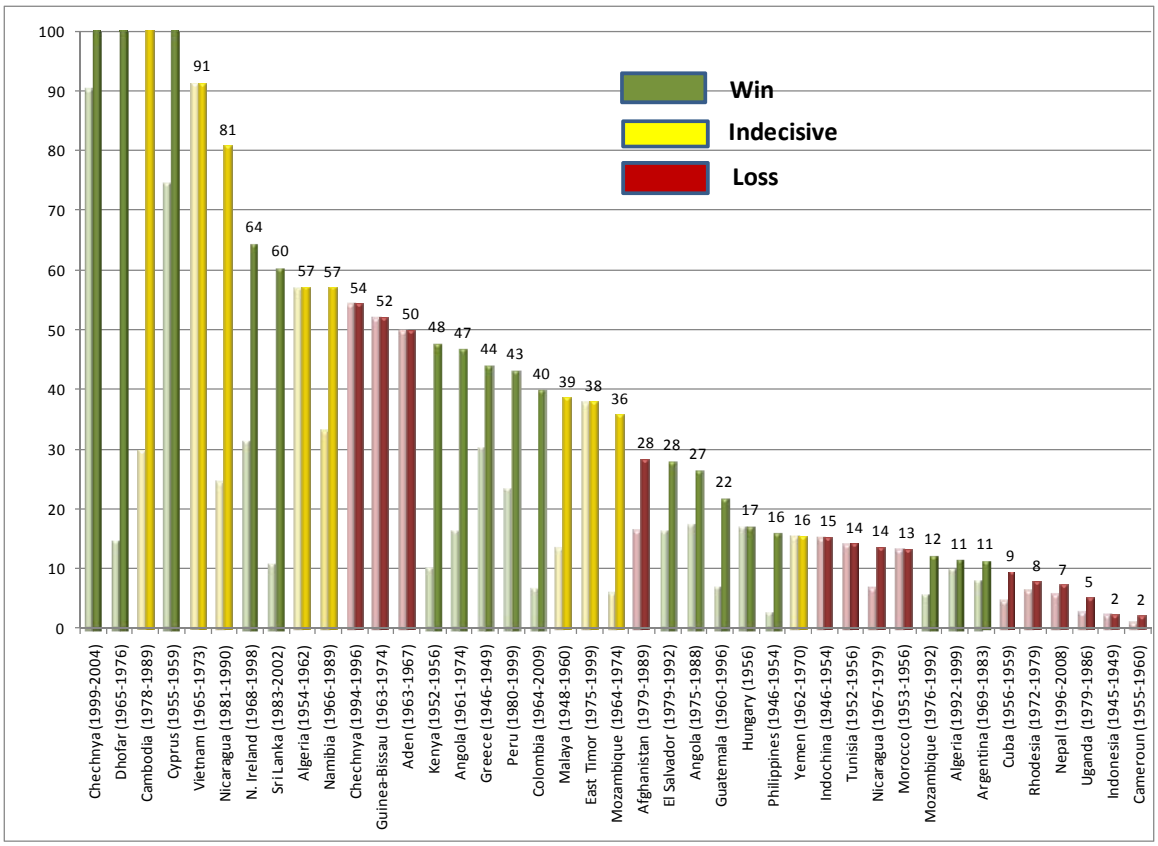


Figure 1. Selected Historical COIN Operations with Peak Force Densities

Figure 1 displays information for the historical COIN cases that were selected. The ordinal axis shows force densities. The darker bars show the peak force densities achieved in the *area of operations* for each conflict[4] and are color coded using the study's scoring of the outcome of the conflicts. The lighter bars show the force densities for the entire country (i.e., COIN troops divided by the country's population). The graph shows that force densities in the actual areas of operations correlate better with outcomes than force densities for the entire country—consistent with FM 3-24 guidance that the force density factor should be applied to the area of actual operations. (Most other analyses of such data do not make the reductions to the actual area of operations.)

In addition to reducing populations to the areas of actual operations, we discovered that the definition used for operational "success" was important in deriving a relationship between force densities and outcomes. We used 'win" or "indecisive" as our

4 We scoped each of the historical cases to the area of operations wherein COIN missions were conducted. This provided a more accurate scaling of population and geography since, in most cases; the actual operations were confined to distinct regions of a country.

definition of success, since in these two situations, the COIN force was able to prevent a military win for the insurgency

Our analysis of these data led to the following findings:

- There is validity to the guidance offered in FM 3-24 that a force density of 20 troops per 1,000 in the area of operations is the *minimum* required for effective COIN operations; however, force densities at that level carry significant risks.
  - Of the 41 conflicts in our data set, the 25 cases with force densities greater than 20 show four losses (16%), while the 17 cases with force densities below 20 show 10 losses (62.5%). For the 14 cases with force densities less than 15, 71% were losses.
- Force densities on the order of 40 per 1,000 inhabitants provide a significantly higher likelihood of success. Thus the suggestion in FM 3-24 that 20-25 troops per 1,000 is an appropriate range of force densities for COIN operations was *not* confirmed by the study.
  - For the eight cases with force densities in the 20-40 range, only 50% were deemed successful. For the 18 conflicts with force densities 40 or greater, 15 were deemed successful (83%).
- Three conflict cases with force densities greater than 50 that were scored as losses demonstrate that strategy, tactics, and non-military/political factors are at least as important as force density in determining success in COIN operations (and, the force size must be appropriate for successful implementation of strategy and tactics).
- Statistical analyses using logistic regression techniques strengthen support for the above findings:
  - Force density is *statistically significant* in predicting successful outcomes for COIN operations at the 5% confidence level (see Figure 2). We also found statistical significance for a binary independent variable defined by whether the COIN force largely comprised forces from developed nations. This result suggests that a COIN force comprising mostly Western forces will have greater difficulty in achieving success than one largely comprising indigenous forces.

Variable	Blue Wins	Blue Does Not Lose
Force Density (Adjusted)	.251	**.011**
Force Ratio	.080	.311
Number Red Factions	.977	.833
Motivation = Nationalism	.091	.758
Blue = Developed Nation	.126	**.028**

Figure 2. Statistical Analysis of Historical COIN Operations

- The logistic regression allows construction of a distribution for probability of success versus force density (see Figure 3). A force density of about 16 corresponds to a probability of success of 50%, whereas a force density of about 40 corresponds to a probability of success of 75%.

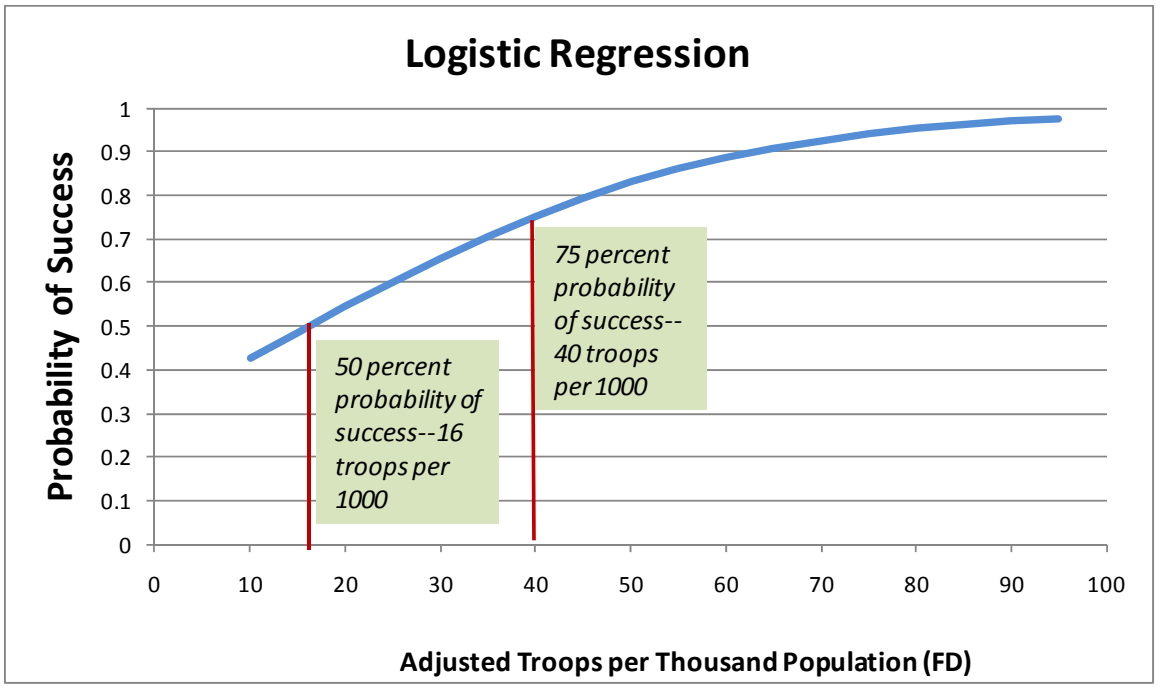


Figure 3. Logistic Regression Results

These findings depend heavily on three factors: (1) reducing country populations to the *actual areas of operation* (see details below), (2) defining "success" as noted above—i.e., the COIN force did not suffer a military defeat (even though the insurgency

may have ultimately prevailed for political reasons), and (3) scoring several historical cases that other researchers (using different criteria) coded as "failures" as "indecisive." Changing any of these factors causes the statistical significance to fail.

3. Findings from Other Research Organizations

We did not find any other study that identifies a statistically significant relationship between force density and COIN campaign outcomes. Two other research organizations[5] tested for such a relationship, also using logistic regression analysis of historical data, but did not find force density to be statistically significant. A comparison of the data and methods employed in those studies points to differences in the three factors identified in the paragraph above. In addition, there were differences in the data on the number of troops actually employed in certain campaigns. Only *one other study*[6] performed the analysis for the actual areas of operations, and for that study there were differences in defining the boundaries of the conflict area. The most important factor is the differences in scoring the outcomes. (There is no clear standard for categorizing a given outcome as a success, a failure, or indecisive. We compared the outcome scorings for different researchers and found a consensus in only about one-third of cases that were in common.)

4. Insights from Ongoing Operations

Recent operations in Iraq and Afghanistan offer a rich source of valuable data on COIN-type stability operations. This study's analysis of that data was necessarily limited by the resources available and access to data. Nonetheless, several useful insights could be drawn. First, we found additional corroboration with regard to the relationship between force density and campaign success. Second, we found data to support an assumption about the mix of indigenous and foreign intervention counterinsurgency forces, which was needed to complete our estimate of U.S. force sizes for success in stability operations. Specific findings include that:

- The level of troops reached in Iraq for the summer 2007 surge, counting U.S., allied and partner nations, and indigenous Iraqi security forces (military and

---

5 CAA and the Defence Science and Technology Laboratory (DSTL) of the United Kingdom Ministry of Defence.

6 By DTSL. See Andrew Hossack and Karthik Sivasankaran, "Success Factors in CT/COIN Campaigns: Preliminary results arising from current research" (Presentation, Cornwallis Group X: Analysis for New and Emerging Societal Conflicts. The Canadian Royal Military College, Kingston, Ontario, Canada, 21-24 March 2005).

police) provided an overall force density of almost exactly 20 per 1,000 when counting the entire Iraqi population. The ratio for the actual area of operations could not be determined but would be significantly higher.

- The ratio in Iraq of U.S. and coalition forces to indigenous forces was in the range of 1:1 to 1:3
- By contrast, in Afghanistan, the force density in July 2007 was only about 5 per 1,000 (although it is considerably higher now).

5. Force Size Projections

The study used a three-step process to develop its estimates for the sizes of U.S. forces needed for success in future large-scale (COIN-like) stability operations in a wide range of countries using a population-centric strategy. The steps are

1. Estimate the total "Blue" force size needed (including U.S., coalition, and indigenous forces)
2. Estimate the proportion of the total Blue force comprising foreign intervention forces, and
3. Estimate the proportion of the intervention force comprising U.S. forces.

For step 1, three different approaches were used to project the total Blue force. The first approach applies the likelihood curve (Figure 3) derived from our logistic regression analysis. Two force densities were chosen, corresponding to probabilities of success of 50% and 75%. Because of concerns about the size and quality of the data supporting the logistic regression analysis, a second approach was developed using a linear regression on the total Blue force size (dependent variable) and the population of the area of operations (independent variable) for 30 conflicts with a Blue success (as defined above). For the purpose of comparison, we also prepared a set of projections based on the updated 2005 methodology (using Iraq as the reference operation). In almost every case, the 2005 methodology yielded a lower force size projection for two reasons: the force density for Iraq was somewhat below the coefficients derived from historical data, and the inclusion of factors other than population tempered the results.

For step 2 we selected a value of 2:3 for the ratio of Blue intervention to indigenous forces for typical COIN operations, based on data cited above which was drawn from current operations and historical data. For the third step, the fractions of the intervention force comprising U.S. forces for each country were drawn from (updated) subjective factors originally developed for the 2005 analysis. Those factors considered the country's location, historical relationships with partners or allies, and the country's

relative importance to U.S. strategic interests compared to that of allied and partner nations.

The main body of the report provides results for each step in the process and displays the projections for all four estimating approaches. Figure 4 displays the study's "bottom line" estimates for U.S. forces for the 26 largest countries for which evaluations were made. These estimates utilize the linear regression approach, which yields mid-range values among the techniques used.

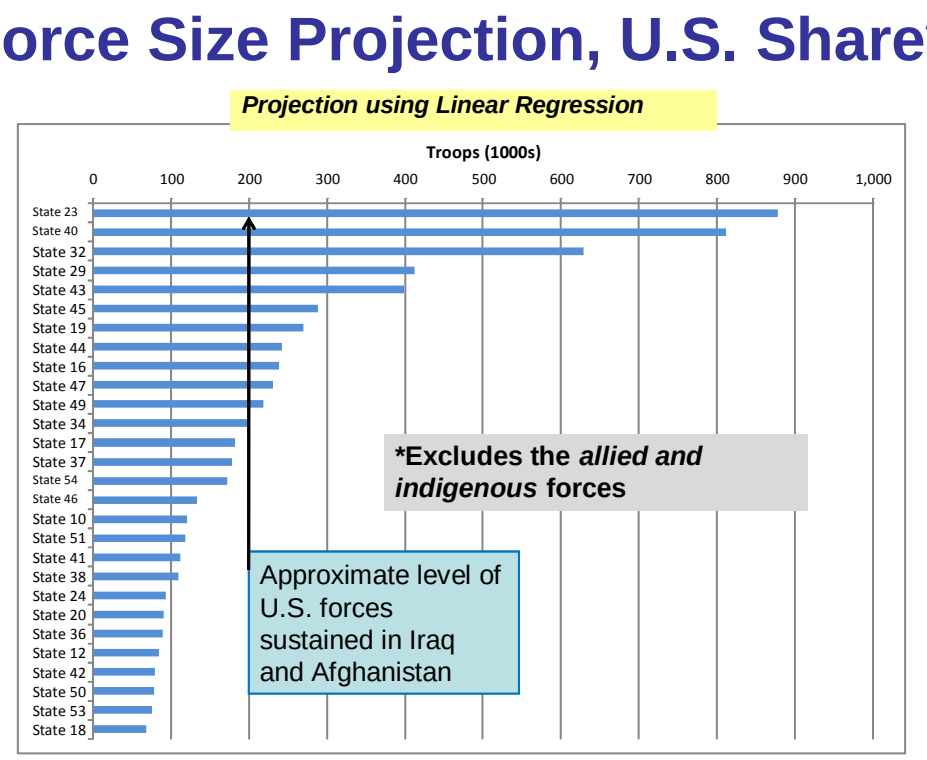


Figure 4. Force Size Projections for Stability Operations

A wide range of unavoidable uncertainty surrounds these estimates—on the order of a factor of two or more. Using mid-range estimates, the current U.S. ground force posture could probably sustain a COIN-like stability operation in most of the countries considered, but several key larger-population countries would likely be infeasible, unless the area of operations within the country could be limited significantly. (See the Classified Appendix for country identities.)

The study's projections for stability operations associated with Defense Planning Scenarios are compared to estimates made within the DoD Analytic Agenda process in the Classified Appendix of this paper. Most of the DoD estimates fall within the range of our projections, but two fell outside the range and thus may deserve closer examination.

C. CONCLUDING REMARK

Projecting force requirements for future stability operations is subject to a large degree of uncertainty. Although the quality of data available on force levels for historical operations has improved significantly over the past few years, we still do not know enough about the past to be highly confident in understanding relationships between force sizes committed and outcomes. Nonetheless, the analysis now possible represents a significant step forward compared to the analysis on which existing doctrine, such as FM 3-24, is based. Differing from other studies, this study found a statistically significant relationship between force density and avoidance of defeat, though it is dependent on key assumptions. Such relationships provide only a starting point for sizing the force for actual operations. Developing the right strategy and tactics is at least as important as force size, and the political landscape is an overarching factor. The key to increasing the level of confidence in our results is better data.

D. RECOMMENDATIONS

- **Revise Army Field Manual 3-24 to reflect more recent analyses of historical data since its publication in 2006.**
- **Conduct additional research of historical data.** Given the sensitivity of our results to factors discussed above, more research into historical data is warranted. For example, selected conflicts could be examined in greater detail to understand exactly what the military parameters and outcomes were over time. Additional conflict cases may now be available that could be included in our sample.
- **Capture, organize, and archive data from ongoing operations.** Although this study was able to draw on the vast amount of data collected from operations in Iraq and Afghanistan, we did not find the data to have been organized, categorized and archived in a manner that facilitates analysis. Those data are an extraordinarily valuable resource for future analysis of stability/COIN operations, but the value will be significantly lessened if the data are not better organized and archived with appropriate controls over access.
- **Define demands for a "whole of government" approach to COIN-like stability operations.** This study, by design, focused solely on military force size; however, meeting demands for a broader range of services, including security, public safety, economic, political, justice, social and public welfare, are of primary importance. Although those activities go well beyond traditional military tasks, experience in recent operations has demonstrated convincingly that there is a strong military role in effectively addressing such demands. A follow-on effort would analyze the appropriate role of the

military in the “whole of government” effort and the relationships between military and civilian activities. A first step would be to develop a “production function” for each required line of effort.

2. ANALYSIS (ANNOTATED BRIEFING)

ORGANIZATION OF THIS SECTION.

This section presents our analysis in the form of an annotated briefing. A page of notes that expand on or offer additional explanations for materials presented on the slide appear above each slide of the briefing. In some cases, there is more explanatory material than will fit on one page of notes. In such cases, this material will be in the notes for the slide immediately following the slide in question. The reader will be cued for such occurrences by the phrase “for [or in] the slide following…”

A Classified Appendix, Appendix D, is published separately and includes several of the displays in this section with the countries identified by name. It also includes a key to the countries referred to by number in this briefing and the application of the force size projection methodologies to selected Defense Planning Scenarios.

A. Tasking, Approach and Context

Stability operations are a broad set of tasks, missions, and activities conducted in concert with all elements of national power in order to maintain or reestablish security and basic services to areas beset by disaster and strife. Department of Defense Instruction (DODI) 3000.05 establishes "stability operations [as] a core U.S. military mission that the Department of Defense shall be prepared to conduct with proficiency equivalent to combat operations."

DODI 3000.05 also directs the Office of Director, Cost Assessment and Program Evaluation (CAPE) to "assess the sufficiency of resources related to irregular warfare and stability operations within the Future Years Defense Program." The Institute for Defense Analyses (IDA) was asked to conduct this study to assist in that task. The purpose of the study is to provide DoD leadership with a rough idea of the capabilities and limitations of the programmed ground force structure to conduct stability operations in a broad range of countries. The estimates of this study should be not used for operational planning (though they may serve as a starting point), rather their intended use is for programmatic force planning and evaluation.

Background for the current study is found in work IDA performed in late 2005. In supporting the 2006 Quadrennial Defense Review (QDR), the then Office of Director, Program Analysis and Evaluation (now CAPE) asked IDA to conduct a quick-turn analysis to produce broad-brush estimates of ground force size requirements for stability operations in selected countries of interest. The objective was to apply some "outside-the-box" thinking about potential force demands for a broad range of stability operations, in contrast conceptually to the much higher resolution analyses of force sufficiency being conducted within the Department's "Operational Availability" series of studies. (Those studies were limited to the few scenarios in the set of approved Defense Planning Scenarios (DPS).)

CAPE asked IDA to update and refine its analysis. Specifically, we were asked to revise the force-sizing model developed in 2005 by reviewing the most current research on stability operations and by incorporating lessons learned from ongoing stability operations.

The Classified Appendix identifies the countries for which force-sizing estimates were developed and provides estimates for the number of forces necessary to conduct the operations specified in the Department's DPSs.

The Tasking and Approach

- Sponsor: OSD(CAPE) (formerly PA&E), Irregular Warfare Div.
- Purpose: Analyses that will contribute to the Quadrennial Defense Review (QDR) and support CAPE Program Reviews
  - Develop force-sizing estimates for programmatic force planning and evaluation **(*not for operational planning)***
  - Underlying issue: How much force would be required to conduct successful stability operation in various countries (as wide a range as possible)?
- Approach
  - Update and extend analysis conducted for the 2006 QDR, supported by IDA, that developed a "Multi-Factor Scaling Model" with current data and evaluate its analytical properties
  - Conduct review of literature, historical data, and models
  - Apply lessons learned from current operations
  - Develop parametric models using statistical techniques
  - Develop force size projections for future potential stability operations
  - Apply revised force-sizing methodologies to estimate the force sizes for stability operations explicit or implicit in the Defense Planning Scenarios

The slide below outlines the analytical process that we used in response to our tasking , and also serves as an outline of this briefing.

Preliminary to the analysis, it was important to clarify the definition of stability operations for the study's purpose and to frame the parameters for force-sizing estimation for the intended purpose. In addition, it was important to understand the context and implications of Army Field Manual FM 24-3, which was published after completion of the 2005-2006 study.

As called for in the task order, we updated the factors that were used in the model developed for the 2005-2006 study and, more broadly, re-examined the overall premise and assumptions on which the model was based.

In 2005 there was a paucity of data on historical stability operations, and the data that did exist were from diverse sources, were incomplete, and of limited reliability. Given the short time available (only a few weeks), the 2005-2006 effort drew largely on the operations then ongoing in Iraq, and to a lesser extent, Afghanistan, though data from Bosnia and Kosovo operations were also employed. Fortunately, in the ensuing years, considerable additional research and data collection have been accomplished by several research organizations. In particular, the study was able to capitalize on a major undertaking by the Center for Army Analysis (CAA) to compile an extensive database on over 100 "irregular warfare" conflicts since World War II, completed just in time for our use. From that data set, our study team extracted some 41 conflict cases—all those that met the two criteria of being a large-scale counterinsurgency (COIN) operation and having sufficiently complete data on COIN force sizes.

Using the data extracted and adjusted from the CAA data, we conducted statistical analysis to find statistically significant relationships using the logistic regression technique.

Ongoing operations in Iraq and Afghanistan offer a rich source of valuable data on COIN-type stability operations. The study's analysis of those data was necessarily limited by the resources available and access to data. Nonetheless, several insights useful to the study could be drawn.

Finally, we drew upon our analysis of both historical and ongoing operational data to develop our estimates of force sizing for future stability operations. It is important to note that these estimates were made for programmatic purposes, not for operational planning.

Outline of Analytical Process

- Preliminaries:
  - Review of the definition of “stability operations”
  - Consider the parameters that might have an impact on force sizing for stability operations
  - Review the guidance offered by Army Field Manual 3-24
- Update, review, and evaluate the methodology used in the 2005 IDA study
- Research the historical data on stability operations that have become available since 2005
- Conduct statistical analysis of the selected historical data
- Investigate data from ongoing operations in Iraq and Afghanistan for useful insights
- Apply the analysis to size forces for future stability operations

The slide below presents the definition of “stability operations” as reflected in recently-published Department of Defense Directives. From a force-sizing perspective, the most important element of stability operations as defined by the directives is the need “to maintain or reestablish a safe and secure environment.” The parallelism implied by the phrase “to maintain or reestablish” may not be the best characterization, since *maintaining* a safe and secure environment may entail no more than a small-scale training mission, whereas *reestablishing* a safe and secure environment may require the deployment of hundreds of thousands of troops for a period of several years.

We focused on COIN or COIN-like stability operations because they entail a much greater commitment of forces than other types of stability operations, thus driving the overall demand for stability forces. (By “COIN-like,” we mean operations for which there is a definite possibility of significant combat.)

The *slide following* discusses the range of stability operations that the study addresses. In practice, COIN operations occupy the high end of this spectrum, whereas other stability operations are likely to fall at the lower or lower-middle end. This study makes the *assumption* that force demands for stability operations that entail COIN operations within the next decade drive programmatic force demands for *all* larger-scale, mid-intensity stability operations. Thus, IDA’s analysis focuses on COIN-related demands. Other large-scale stability operations that aren’t really COIN operations may be sufficiently like them to have similar force demands. Examples would be the Bosnia and Kosovo operations of the late 1990s. Bosnia turned out to be a large-scale peacekeeping operation rather than COIN-like combat, but there was clearly a threat that the warring forces (especially the Serbs) would turn to insurgency tactics. Similarly, Kosovo also was largely peacekeeping, but there was a clear threat that combat might ensue with Serbian forces. Thus, arguably, these conflicts fall within the range of larger-scale, mid-intensity stability operations for which force sizing based on historical COIN operations makes sense.

What Are Stability Operations?

An overarching term encompassing various military missions, tasks, and activities conducted outside the United States in coordination with other instruments of national power **to *maintain* or *reestablish* a safe and secure environment, provide essential governmental services, emergency infrastructure reconstruction, and humanitarian relief.**

DoD Directive 3000.07, Irregular Warfare

"Stability operations are a core U.S. military mission that the Department of Defense shall be prepared to conduct and support. **They shall be given priority comparable to combat operations** . . ."

DOD Directive 3000.05, Military Support for Stability, Security, Transition, and Reconstruction

In order to size the force for an *actual operation*, it is essential to conduct a troop-to-task analysis that fully accounts for situation-specific details, such as the precise nature of the mission and the contours of the human and geographic terrain. In contrast, sizing decisions for the programmed force depend on the estimated demand associated with a broad range of future operations, *whose nature cannot be fully known in advance.* Certain basic parameters of the area of operations (AO) for those future operations, however, may be known, or reasonably postulated. Such parameters include the size and density of the population, the extent of urbanization, the size of the AO, and similar factors. Important factors, such as the size of an insurgency force, are subject to far greater uncertainty. Thus, a parametric approach to estimating force requirements is the most viable for program planning and evaluation. Knowing this much only defines the problem to be solved, however. The relative importance of different parameters has been the subject of minimal research. An objective of this study was to assess the significance of such parameters.

Although estimating the size of the force is essential for planning purposes, COIN experts continually emphasize that success or failure depends on the choice of appropriate strategy and tactics, as well as non-military factors. Classic texts on COIN, such as the writings of David Galula, Roger Trinquier, and Sir Robert Thompson, pay scant attention to the question of force sizing. Instead, they assert time and again that success depends on securing the population, more than anything else. Authors in the present day also emphasize the decisive role of strategy and tactics, as well as non-military factors. Yet presumably, there is a minimum number of personnel required to implement even the best strategy and tactics. Moreover, assuming that one's strategy and tactics are sound, does the introduction of additional forces improve one's chances of success? If so, is there a point of diminishing returns at which the introduction of too many personnel antagonizes the local population? Is the "Powell doctrine" of overwhelming force applicable to stability operations, and if so how? (An issue raised in a recent speech by the current Chairman of the Joint Staff.) Finally, are sufficient data available, in terms of both quality and quantity, to answer these kinds of questions? This study will argue that, although more and better data remain extremely desirable, the data now available can be used to define a range of force-sizing estimates which, though broad, is reasonably well-supported analytically.

Key Framing Issues

- Stability Operations span a diverse range of operations with different drivers of force requirements (e.g. COIN, Peacekeeping, HA/DR*)
  - Operations expected to involve significant combat (e.g., COIN) demand the largest number of forces
  - Less demanding operations, limited in frequency and duration, do not drive programmed force size
  - ***Thus, this study focused on COIN***
- Programmatic force planning is best served by a parametric approach, while operational planning requires "troop-to-task" analysis (requiring much more situation-dependent information)
- This analysis attempts to bound a reasonable range of force size requirements using historical lessons
  - *Strategy and tactics are at least as important as force size in determining success*
- Success in stability operations depends heavily on non-military factors (but outside our scope)

** Humanitarian Assistance and Disaster Relief*

COIN operations comprise several lines of effort– protecting the population, defeating hardcore insurgents, developing the economy, building political and social institutions, etc. Although those activities go well beyond traditional military tasks, experience in recent operations has demonstrated convincingly that there is a strong military role in effectively addressing such demands. Ideally, one could develop a "production function" for each such line of effort. This study, however, focused solely on force size, basically driven by protection of the population.

A number of factors have been suggested as being of significance in force sizing for stability operations. The slide below lists some of these and offers considerations regarding their predictability and potential impact. It would be desirable to be able to relate these or other similar factors to outcomes based on statistical analysis of historical operations; however, efforts in that direction by both the current study team and by other researchers have shown that the ability to establish such relationships based on the available data is rather limited.

This study's examination of population as a predictor of necessary force size will be presented in detail later. (We were, subject to significant caveats, able to establish a statistically significant relationship between the ratio of troops to inhabitants within the AO and conflict outcomes.) The size of AO is also plausibly a factor of significance. We attempted to find such a relationship, but were unsuccessful.

We also examined briefly the degree of urbanization, but discovered that there is no clear relationship between level of urbanization within the AO and required force size. To the extent that an urban area can be cordoned off and entry and exit access controlled, it may be easier to control than an rural area. On the other hand, if troops have to go "house-to-house" to root out insurgents, an urban area may be more difficult to control than rural areas (depending, of course, on the terrain itself).

Later, there will be additional discussion of the relationship between the number of insurgents and force-sizing decisions. For the moment, it should be noted that the number of insurgents cannot be reliably predicted for future operations, thus diminishing its viability as a force-sizing parameter. (Indeed, the ability to estimate insurgent strength accurately, even for historical and ongoing operations, has proven to be notoriously poor.)

Intervention Force Size Drivers for Stability Operations

There are numerous factors that could drive the size of an intervention force required for success in a postulated future stability. Some cannot be estimated with any reliability in advance. A few more salient ones:

Factor	Likely Impact	Predictable
Area/sub-population of a stability operation within a selected country	High	Somewhat
Population	High	Yes
Size	Medium	Yes
Length of uncontrolled border/ coastline	Medium	Yes
Degree of urbanization or Population Density	Uncertain	Yes
Critical infrastructure needing protection	Medium	Yes
Size, loyalty, and competence of indigenous forces	High	No
Size and effectiveness of threat forces	Medium	No

Note: All factors (except the first) apply to the projected area of operations

A key objective of the study was to evaluate and expand on the research that led to the force density suggestions in FM 3-24 as quoted in the slide below.

While asserting clearly that force density (i.e., the ratio of troops to population) is superior to force ratio (i.e., the ratio of troops to insurgents) as a gauge of force requirements, FM 3-24 only cites the opinion of unnamed researchers with regard to what force density is desirable. It is interesting that the FM does not specify a force density rule of thumb as doctrine—assuming that it does is a common mistake found in open literature articles on the subject.

In light of this confusion, the study team contacted the authors of the FM to make sure of their intent. They responded that it was necessary to establish some benchmark for the appropriate force density, but the research underlying the figure of 20-25 troops per 1,000 inhabitants was too uncertain to justify a more explicit endorsement. They ultimately wrote:

Paragraph. 1-67. The movement leaders provide the organizational and managerial skills needed to transform mobilized individuals and communities into an effective force for armed political action. The result is a contest of resource mobilization and force deployment. No force level guarantees victory for either side. During previous conflicts, planners assumed that combatants required a 10 or 15 to 1 advantage over insurgents to win. However, no predetermined, fixed ratio of friendly troops to enemy combatants ensures success in COIN. The conditions of the operational environment and the approaches insurgents use vary too widely. A better force requirement gauge is troop density, the ratio of security forces (including the host nation's military and police forces as well as foreign counterinsurgents) to inhabitants. Most density recommendations fall within a range of 20 to 25 counterinsurgents for every 1,000 residents in an AO. Twenty counterinsurgents per 1,000 residents is often considered the minimum troop density required for effective COIN operations; however as with any fixed ratio, such calculations remain very dependent upon the situation.

Paragraph 1-68. As in any conflict, the size of the force needed to defeat an insurgency depends on the situation. However, COIN is manpower intensive because counterinsurgents must maintain widespread order and security. Moreover, counterinsurgents typically have to adopt different approaches to address each element of the insurgency. For example, auxiliaries might be co-opted by economic or political reforms, while fanatic combatants will most likely have to be killed or captured.

Army Field Manual 3-24 (2006)

Counterinsurgency

"Most [force] density recommendations fall within a range of *20 to 25 counterinsurgents for every 1,000 residents in an AO [area of operations]…*Twenty counterinsurgents per 1,000 residents is often considered the *minimum troop density* required for effective COIN operations; however as with any fixed ratio, such calculations remain very dependent upon the situation." *(Paragraph 1-67)*

This study evaluated and expanded on this suggested rule of thumb

FM 3-24 is both extremely influential and widely misunderstood. Its rule of thumb about force sizing is frequently cited by journalists, politicians, and prominent analysts. Most accounts suggest that FM 3-24 endorses the target of 20 troops per 1,000 inhabitants as the baseline figure for force density in COIN operations. The reality is much more complicated. First of all, FM 3-24 clearly asserts that force density is superior to force ratio as an indicator for determining how many boots to put on the ground. Next, FM 3-24 observes that other, unnamed analysts consider 20-25 troops per 1,000 to be the appropriate range. Finally, it adds that unnamed researchers consider 20 per 1,000 to be the minimum force density required for effective operations.

Almost all references to FM 3-24 blur the distinction between these last two observations. Instead of explaining that there is one recommended minimum (20 per 1,000) and a different recommended range (20-25 per 1,000), they simply report a single recommendation of either 20 or 20-25 per 1,000. Analytically, these propositions are quite different. The notion of minimum suggests that densities below 20 per 1,000 involve a considerable risk of mission failure. The notion of range suggests that, whereas 20 per 1,000 is the effective minimum, there is little reason to exceed the figure of 25 troops per 1,000 inhabitants. This study's analysis of 41 historical COIN campaigns (described in detail later in this brief) indicates that while the minimum recommended density is 20 per 1,000, successful campaigns have often involved ratios of 30-50 per 1,000 – and sometimes much higher.

It should also be noted that few analysts have rigorously studied the relationship between force density and success in COIN operations. James Quinlivan's seminal work on this subject did arrive at the figure of 20 per 1,000, although it considered only a handful of operations. Both John McGrath and the team led by Amb. James Dobbins arrived at recommendations in the low teens, although they also considered a limited number of operations, which included several that weren't COIN. In effect, there is, at best, a very limited consensus on the appropriate force density for stability and/or COIN operations.(See Appendix C for full citations.)

Understanding FM 3-24

- **FM 3-24 makes three analytically distinct assertions:**
  - **Force density is superior to force ratio as an indicator of how many boots to put on the ground**
  - **The *minimum* recommended force density is 20 troops per 1,000 inhabitants in the AO**
  - **The *range* of recommended force densities is between 20 and 25 troops per 1,000 inhabitants in the AO**
- **The findings of this study support the first two statements, but not the third**
  - **The observed force densities in relatively successful campaigns were often 30-50 troops per 1,000 inhabitants and sometimes even higher**
  - **A force density of 20 per 1,000 is plausible as a minimum, but our research does *not* confirm 25 as an upper limit (although it is not clear that was the intention, but it is a possible interpretation)**

The slide below presents an overview of the three separate methodologies that were used to estimate force sizes for future stability operations. These approaches will be explained in more detail subsequently.

The first method is the subject of the next section of the briefing. Both the second and the third methods are based on historical data and, of course, their validity depends heavily on the quality of that data.

The second method, using logistic regression, employs data based on 41 historical contingencies that have been characterized as "successes" or "failures." For this analysis we have defined "success" as either a "win" or "no lose." The latter can be characterized as "military successes" in that the military operations prevented an insurgent military victory, even though there may have been subsequent political concessions with ambiguous outcomes. (This characterization is *key* to the study's findings and will be addressed later in the briefing and discussed more extensively in Appendix, A wherein the scoring for each of the 41 cases is discussed.)

The third technique is based on a linear regression using 25 of the 41 contingencies historical campaigns that were "successes," plus 3 more recent campaigns (Bosnia, Kosovo, and Iraq) that are also characterized as such.

As noted, all the techniques provide an estimate of the total size of the required "Blue" force, which would comprise U.S., allied or partner, and indigenous forces. So, once an estimate of the total Blue force is obtained, another set of factors must be applied to determine the U.S. component of that force. We will explain how such factors were determined after we have presented the methodologies and results for the first step.

Note: Some analysts refer to allied and indigenous forces as "Green" rather than Blue. We use "Blue" to refer to all COIN forces. When appropriate we differentiate the components of Blue forces.

Alternative Methodologies for Projecting Force Size for Stability Operations

Multi-factor scaling methodology	**The 2006 methodology (scaled from Operation Iraqi Freedom (OIF))**
Troop density (troops per 1,000 population) multiplier derived from logistic regression analysis	**Based on probability of "success" (50% & 75%) from a logistic regression analysis of historical data from 41 conflicts**
Linear regression equations	**Regression equations based on 28 selected historical operations categorized as "success"**

These methods provide *total troop requirements*, which can be obtained by a mix of intervention (U.S. and allies) and competent indigenous forces. Thus to obtain U.S. troop requirements, it is necessary to address this force mix issue—i.e. the required ratio of intervention (U.S. and allies/partners) to indigenous forces

The discussion so far sets the stage for our analytical results, findings, and conclusions, which will be presented as indicated on the slide below.

Presentation of Analytical Results

- The 2005 multi-factor scaling methodology
- Determining actual or projected area of operations
- Analysis of historical data on large-scale stability operations
  - The dataset and its characteristics
  - Statistical analysis using logistic regression
  - Comparison to findings of other research organizations
  - The linear regression approach
- Comparison of force size projection methodologies
- Projecting U.S. force requirements for stability operations
  - Insights from recent COIN operations in Iraq and Afghanistan
  - Ratio of intervention to indigenous forces
  - Level of allied/partner participation
- Force size projections for stability operations
  - Diverse range of 54 countries

B. The 2005 Multi-Factor Scaling Methodology

We begin by discussing characteristics, advantages, and drawbacks of the methodology used in the IDA study on this topic that was conducted in 2005-6.

Examination of this topic for the 2006 QDR led to the construction of a simple and transparent Multi-Factor Scaling Methodology (MFSM) for estimating the number of troops necessary to conduct stability operations. In order to estimate the size of the force required to conduct operations in a specific country, the MFSM scales the force size from a historical (reference) operation, based on six parameters. They include the size of the country, the size of its population, its stability based on the Failed States Index (FSI), a subjective "receptivity/complexity" factor, the presence or absence of weapons of mass destruction (WMD), and the estimated percentage of combat power to be supplied by other members of the intervention force.

The technique is simple. For population, area, and the FSI, compute the ratio of those parameters for the country of interest to those for the reference operation. The other two factors are pure multipliers (not ratios to the reference operation). Thus, to derive an estimate for a stability operation in Country A, one would form the ratios of Country A's size, population and FSI score to those parameters for Iraq. The product of those three ratios is then multiplied by the other two factors , and then by the size of the force needed in Iraq to obtain an estimate of the size of the intervention force. Finally, that result may be multiplied by a factor reflecting the expected level of allied and partner nation participation to obtain the U.S. troop estimate. In this example, Iraq serves as the reference operation for estimating the size of the force necessary to stabilize Country A. In order to generate a range of estimates, we employed several different reference operations, including Iraq, Afghanistan, and Bosnia.

Illustrative calculation:

	Population (M)	Area (MSqKm)	Failed State Index	Popu-lation ratio	Area ratio	Stabiltiy ratio	Receptivity-'Complexity	WMD Factor	Inter-vention Troops Required	Level of Allied part	U.S. Troops Required	
Country A	39.1	0.15	85	1.48	0.68	0.77	0.9	1	**279,852**	0.15	**237,874**	*Computed*
Reference Operation	26.5	0.22	110						400,000			*Given*

2006 Multi-Factor Scaling Formula

- Scaling factor from a historical reference operation (e.g., Iraq)
  - { Area ÷ (Area of Iraq)×0.25 + Pop. ÷ (pop. of Iraq)×0.75 }

    ×
  - Stability Index relative to Iraq

    ×
  - Receptivity/Complexity Factor (subjective)

    ×
  - Weapons of mass destruction (WMD) factor (subjective)
- Multiply by the *intervention force size* of the reference operation to obtain the required size of the intervention force
- Multiply by the assumed fraction of the intervention force filled by allied and partner nations

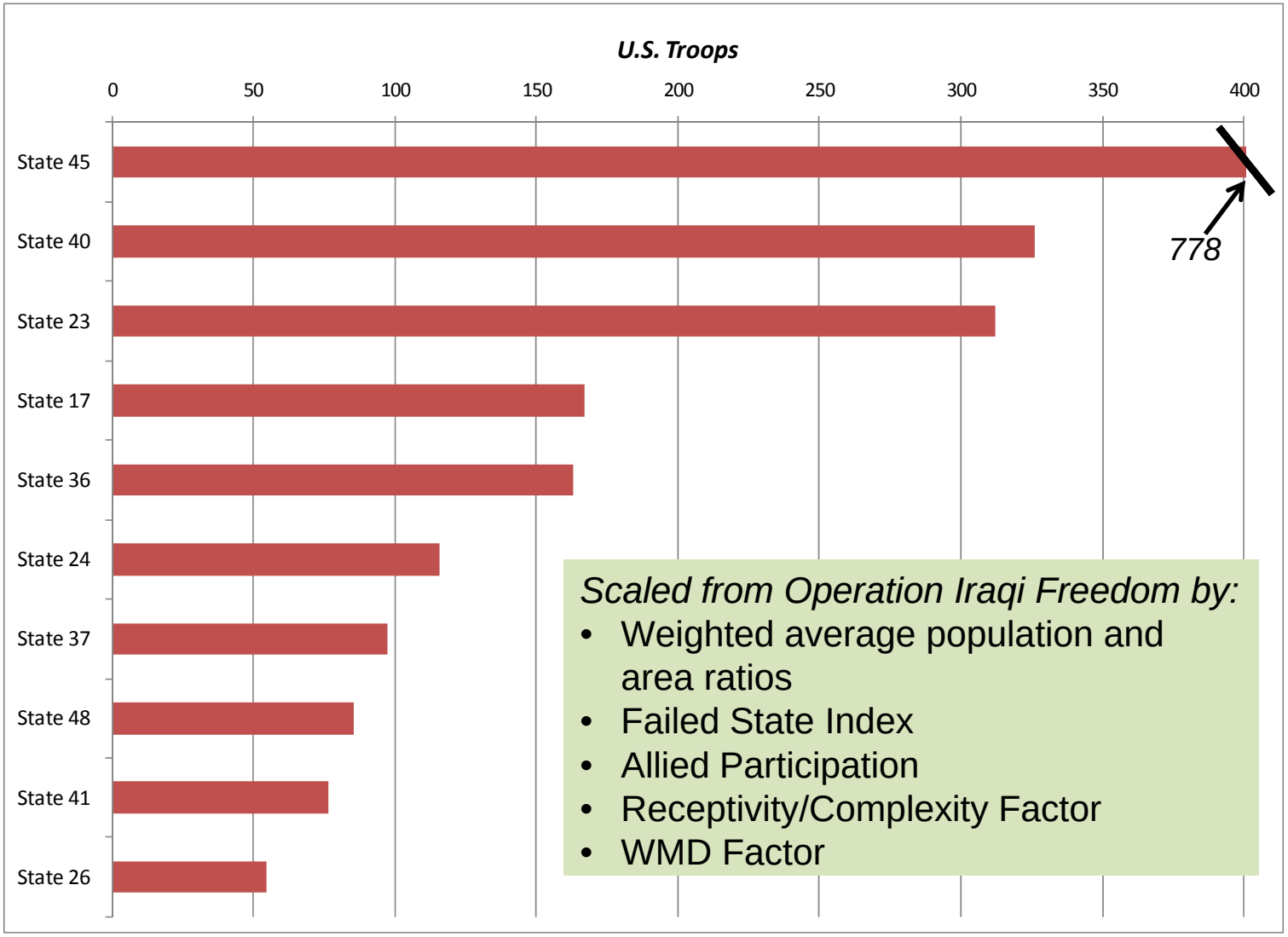
Both the receptivity/complexity and WMD factors were subjective. The WMD factor was set at one if the country was not believed to have WMD that would require securing. If possession of WMD was believed or postulated for the target year, it was set to 1.25 or 1.50 depending on the number of suspected sites. The receptivity/complexity factor was set based on an assumption about how receptive the population might be to a foreign intervention, or it could reflect cultural complexities, such as internal ethnic conflicts.

This slide illustrates results using the MFSM for several countries, the names of which are provided in the Classified Appendix.

These results can be compared to those obtained in 2006 (*Stability Operations: A Range of Possibilities (U),* Kneece, et al, IDA Document D-3231, January 2006). A comparison is provided in the Classified Appendix, including reasons for the differences.

For the current study, the population, area, and stability indexes for each country were updated, as were the parameters for Iraq and Afghanistan as reference operations. For most countries, the difference between the 2006 force-sizing projection and the 2009 projection were minor.

Results from Using the 2006 Methodology
(updated parameters)



This slide illustrates two points regarding the MFSM—the effect of the ratio and factor multiplier calculations and the sensitivity to the reference operation used.

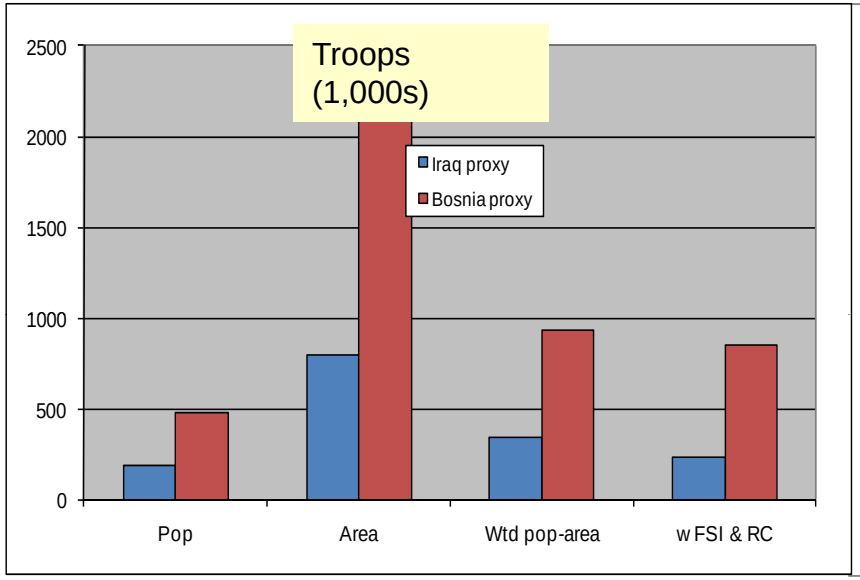
These illustrative calculations are for a medium-size country. Bosnia produces a much higher projection than does Iraq. In 2005, Afghanistan was also used as a reference operation, based on the assessment at the time that the country had been stabilized. Projections based on US/NATO troop strength in Afghanistan circa 2005 yielded results almost an order of magnitude lower than high-end projections based on Bosnia. (Subsequent events showed that the apparent stability in Afghanistan was ephemeral.) The current situation in Afghanistan is not usable as a reference operation because stability has not yet been obtained.

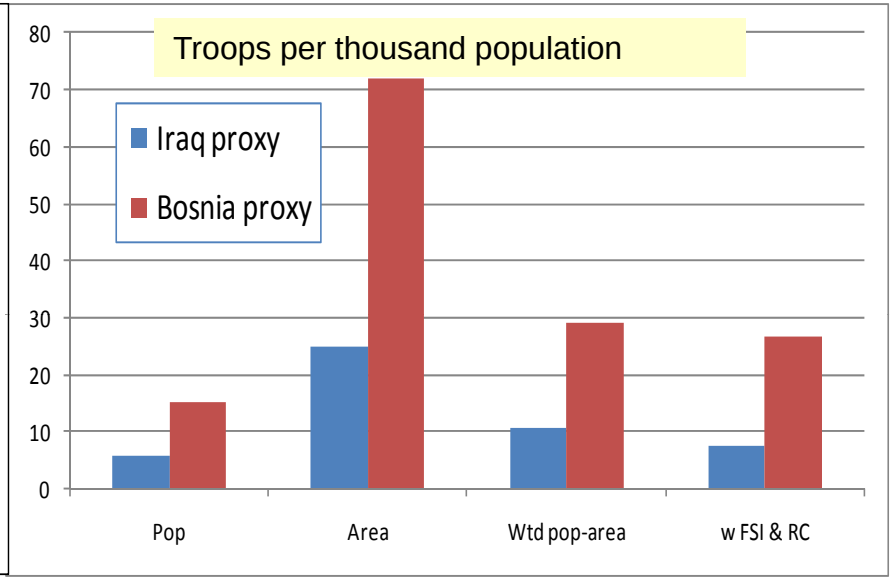
In addition to addressing overall sensitivity, the chart shows the sensitivity of projection to different components of the MFSM. The first set of bars is the results with scaling based on population only, the second set uses scaling based on area only, the third by the weighted average of population and area ratios, while the last shows the impact of adding the stability and receptivity-complexity factors (the WMD factor was 1).

As indicated, we concluded Iraq probably provides the best reference operation for most larger-scale future stability operations in which the U.S. would likely be involved in the next decade.

Illustration of the MFSM Calculation

Effect of Component Factors and Impact of Reference Operation for a Medium-Size Country





- Bosnia and OIF very different operations
- Using Bosnia as the proxy operation results in troop projection about 4x that obtained from using Iraq (OIF)
- Scaling only by weighted population and area results in about a 3x increase
- Scaling from Afghanistan, on the other hand, provides results in the other direction—only 17% of the projection based on OIF

Scaling from OIF provides more credible projections for large-scale future contingencies of the most interest

This slide and the next display the issues and concerns identified by the study team regarding the Multi-Factor Scaling Methodology.

To assess the relevance of area (km^2 in the AO) to force-sizing projections, we looked for correlations between population density and campaign outcomes. It is not clear, *a priori,* whether population density should have a positive or negative relationship to successful COIN outcomes. Whereas some authorities insist that densely populated urban terrain is ideal for insurgents, others assert that this terrain favors COIN forces. Regardless, no correlation emerged from the data even though we ran regressions on several different samples.

We have continued to use the Fund For Peace Failed State Index as a measure of stability in the MFSM. At the beginning of this study, we investigated the available of other stability indices to see if a more appropriate one was available. We concluded that the FSI was still the best choice; however, we observed that none of these indices were *designed* to inform the question of the force size needed for a stability operation, all have some level of subjectivity, and some lack transparency. Even if one accepts the hypothesis that scaling the force size based on a ratio of their stability indices is appropriate, a further complication is that the index that is observed today is not a good predictor for the level of stability that would exist when conditions requiring an intervention occur. For these reasons, the study concluded that use of a stability index to scale the size of a stabilization force is not analytically sound.

As regards the receptivity-complexity factor, we are concerned from an analytical perspective by the entirely subjective nature of the parameter. In addition, it is difficult if not impossible to dissociate that factor and the stability index, since ethnic strife is one of the factors considered in formulating the FSI.

As noted on the slide, the 2006 methodology did not consider the role of indigenous forces. At the time, indigenous forces were not playing a significant role in Iraq or Afghanistan, nor had they in Bosnia/Kosovo. Our perspective on the role of indigenous forces is different today, as will be discussed later. The MFSM can easily be adjusted to generate projections that include indigenous forces.

Issues with Multi-Factor Scaling Methodology (1 of 2)

- Is a weighted sum of population and area scalers better than just population?
  - Intuitively credible but historical data examined does *not* so indicate
- Use of Failed State Index as a scaler
  - Not designed for the purpose
  - Assumption of linear scaling based on "proxy" operation
  - Index will be different (presumably higher) at the time a stability operation is required
- "Receptivity/Complexity" (or "degree of cooperation") factor
  - Purely subjective
  - Some redundancy with use of FSI?
  - Alternative (not pursued): Change definition to "receptivity" only and base on polling data (However, polling data is subject to high variability)
- 2006 methodology did not consider indigenous forces (corrected in 2009)
- Reference operation used to scale may not be representative of a particular future operation (although it may be more representative than many other historical conflicts)

Soon after the start of this study, the team realized that it may be inappropriate to base the size of a stability force on parameters of a country in its entirety, since historical experience shows that frequently only parts of the subject country are in need of a stabilization force. This realization led us to introduce a "reduction methodology," by which the region(s) of a country most likely to be in need of a stability operation would be identified and its parameters obtained. That step was also consistent with the FM 3-24 suggestion that the required force density (troops per thousand of population) should be achieved "within the area of operations." We will report on how these adjustments were made subsequently.

To address the MFSM's shortcomings, we considered the inclusion of new parameters. However, our statistical analysis of historical data did not identify any variables other than population as significant. For certain parameters, such as the length of uncontrolled borders and the amount of critical infrastructure requiring protection, no reliable data were available. Such factors would likely be important considerations, however, in sizing a force for a particular stability operation based on "troop to task" analysis.

Another objection to the MFSM is its dependence on a single reference operation. That criticism, however, is not basic, because it is relatively straight-forward to expand the number of reference operations to as many as desired. The limitation is on the availability of data. While population and area for historical operations are fairly easy to obtain, it is not possible to know, for example, what the FSI might have been for the conditions that prevailed for the historical operation. For that and other reasons, we did not pursue that approach beyond a trial calculation. However, more fundamentally, it can be argued that this "weakness" is in fact a strength. One can carefully choose a well-understood reference operation that is *believed* to be representative of a future conflict involving U.S. forces. The risk of course is that such a belief may be mistaken.

Issues with Multi-Factor Scaling Methodology (2 of 2)

- Not an analytically rigorous approach
- Adjusting the parameters for the target countries to the sub-regions and sub-populations that would be the most likely to require intervention forces should provide more "credible" estimates *for larger countries* and would be consistent with FM 3-24
  - But adds yet another degree of speculation
- The current study could not establish a strong case for introducing any other new scaling factors to the 2006 methodology
- An alternative (which we examined) is to base the scaling on a much larger set of selected historical operations
  - Since it would be difficult if not impossible to develop credible estimates of stability and receptivity for many of the historical operations in our database, we based the calculation on population and area only

The advantages and disadvantages of the 2006 MFSM are enumerated in the slide below. Most of these were discussed in previous slides.

Our overall conclusion regarding the approach is presented as well. The methodology served a useful purpose in 2006 by displaying the range of potential force sizes that might be needed for a stability operation in a large number of countries of interest. There was no pretention that the results should be considered valid for planning forces for specific conflicts; rather it gave top-level decision makers some broad insights into the limitations of the programmed force to perform a demanding stability operation in larger countries of the world.

Now, however, much additional research has been done and additional and higher quality data collected. After considerable work with the MFSM, we concluded that the more recent research and data provide a firmer analytical basis for projecting force size requirements for stability operations. That, basically, is an analytical judgment of the study team that is subject to debate. Thus, we carry forward calculations of force projections using the MFSM for users of the study who prefer that approach.

Conclusion Regarding the 2006 Multi-Factor Scaling Methodology

Method developed for quick-look study for the 2006 QDR

Advantages

- Simple, transparent
- Considers several (presumably) key variables
- Allows adjustment of demand predictions based on intuitive judgments about the effects of subjective factors
- Uses as a "proxy" operation a well-understood, recent conflict with U.S. forces playing a central role—perhaps a reasonable model for future conflicts that might involve U.S. forces

Disadvantages

- Depends on a single reference case as the basis for scaling predictions
- Assumed effects of factors other than population size is not supported by any statistical evidence, nor have we discovered any other sound analytical basis

Historical data provide a more analytically sound basis for stability force size projections

C. Determining Actual or Projected Area of Operations

This section discusses our methodology to adjust country populations to the actual "area of operations" (as stated in FM 3-24). An understanding of process is important in interpreting later results.

In previous studies, the force density was determined by dividing the number of troops by the population of the entire country for the stability operation. However, most insurgencies affect only a portion of the countries in which they occur, and FM 3-24 specifies that the appropriate force density is that in the actual AO for the conflict. Applying that insight affects the assessment of both historical and projected future operations. For historical operations adjusting populations to the actual area of operations (as best that it can be determined) means that the computed force density will be higher than if the population of the entire country were used. In the case of Iraq, for example, there has been almost no violence in the three northern provinces that comprise the autonomous Kurdish region of the country. Thus, we subtracted the population of those provinces (~ 3 million) from the population for Iraq.

Area reductions were also made for input to the MFSM. For Iraq we reduced the area by removing much of the lightly inhabited areas of western Iraq, as well as the three Kurdish provinces. In applying the reduction technique, we took a conservative approach in that only the territory and population removed from the AO were that clearly outside the radius of conflict, e.g., the Kurdish regions of Iraq. For the Malaya and Northern Ireland conflicts, we subtracted the majority populations (Malay and Protestant, respectively) from the AO because they were effectively immune to insurgent influence.

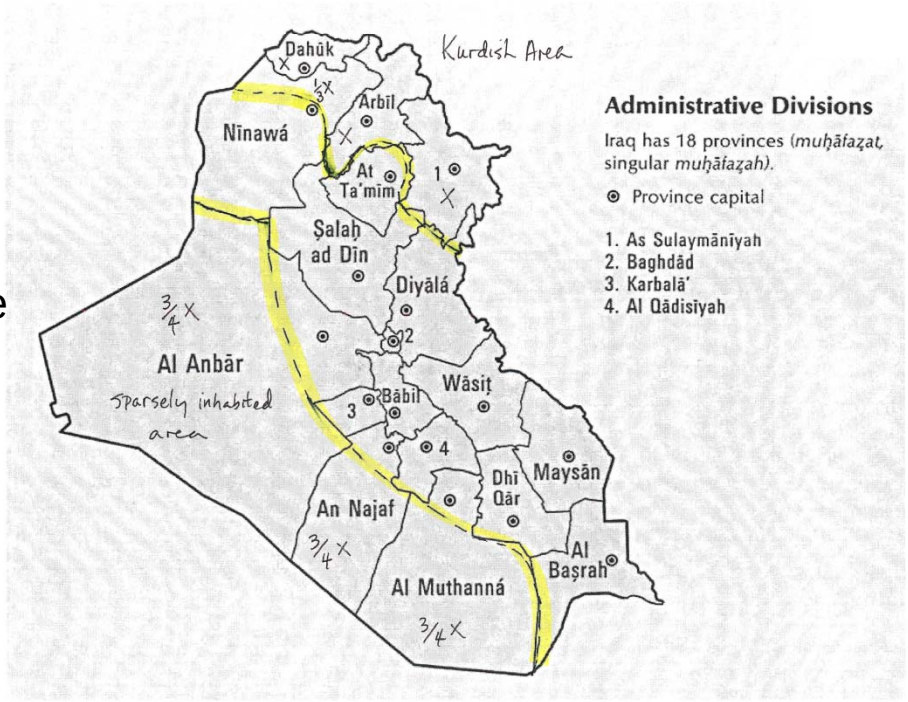
Estimating the area of operations for a projected future conflict is more problematic. In making those adjustments, we considered, based on fairly cursory research, which areas of the subject country would most likely to be where an insurgency might be active. If no such area was readily apparent, we did not make an adjustment. The actual adjustments are displayed in Appendix D.

If one incorporates the adjustment method into the MFSM, it generates lower estimates in the majority of cases. Yet in certain instances, the estimates *increase* because the reduction method also applies to the reference operation. If the percentage reduction for the future operation's AO is less than the percentage reduction for the reference operation's AO, the MFSM estimate will increase. This occurs for all of the countries for which no reduction was made.

Methodology for Adjusting to the "Areas of Operation"

- Consistent with FM 3-24, defines the "areas of operations" within the country subject to the stability operation
- Estimates population and area for the identified AO
- Results in an *increase in troop density* obtained from a historical operation (since the denominator is smaller)
  - The effect on the projected troop size estimate for a target country can be higher or lower
- Easier to apply for historical operations than for projected future operations
- Has a major effect ***(illustration to follow)***

Example (Iraq)



All subsequent analytical results reflect the application of the adjustment methodology

D. Analysis of Historical Data on COIN/Stability Operations

What is the most effective way to analyze data about stability and COIN operations? Our review of the literature indicated that there are two basic strands of analysis. First, there is one set of studies that poses the broad question of what factors influence the outcome of insurgencies and COIN operations. Then there is another set of studies that focuses specifically on the question of force sizing. Both approaches are relevant to our work. As noted in our discussion of the MFSM, the principal challenge facing the designers of a parametric force-sizing model is deciding which parameters to include in the model. If one confirms that a certain parameter has a significant impact on insurgency/COIN outcomes, that is a sound basis for including it in the model. The second step, after identifying the relevant parameters, is to develop an algorithm that translates the value of those parameters into a force-sizing estimate. This section of the briefing presents our statistical analysis of the factors that can be shown to influence COIN/stability-operation outcomes based on historical data.

To conduct sound statistical analysis, it is of course necessary to have sufficient data of reasonable quality. That was not the case for the study done in 2005-2006. Fortunately, since then, there has been a vast improvement in the both the quantity and quality of data available on historical COIN/stability operations. The slide following shows the data sources reviewed and/or used by the current study.

The most important source for our study is an irregular warfare data base commissioned by the Center for Army Analysis, which brings together a wide array of unclassified, publicly available information about 102 different conflicts in the post-World War II era. The CAA Irregular Warfare data base is especially useful for force-sizing analysis, since it contains annualized personnel figures, categorized by national origin and organizational affiliation.

Analysis of Historical Data on COIN/Stability Operations

The main criterion for inclusion in the CAA data set was the availability of sufficient information about a given conflict. Thus, analytically speaking, one cannot assume that the conflicts in the CAA data set comprise a representative sample of irregular conflicts in the post-WWII era. Also, the CAA data set includes all types of irregular conflicts, not just insurgencies. Thus, we sought an independent measure of both the representativeness of the CAA sample and the identification of particular conflicts as insurgencies. With that in mind, we turned to the most comprehensive data set on insurgencies currently available, which was recently assembled by Professor Jason Lyall, Yale University, and Lieutenant Colonel Isaiah Wilson, United States Military Academy.

The Lyall-Wilson data set covers 286 insurgencies since 1815, of which 135 have taken place since WWII. To qualify as an insurgency, a conflict must result in a minimum of 1,000 battle deaths and the challenger to the government's authority must employ the methods of guerrilla warfare. Regrettably, the Lyall-Wilson data set does not contain information on the number of troops committed to COIN operations. Thus, it cannot serve as the basis of a force-sizing analysis.

Of the 102 conflicts in the CAA data set, the Lyall-Wilson data set identifies 57 of them as insurgencies. After excluding those cases for which the CAA data set does not provide sufficient force-sizing data, 41 cases remained suitable for further analysis.

We made two other adjustments to the dataset. First, in spite of sufficient data about Moroccan forces in the Western Sahara campaign, the minimal population of the Sahara made the conflict an extreme outlier in terms of troop-to-population ratio. Thus, we excluded it. Second, we included the case of Hungarian guerrilla resistance to the Soviet invasion of 1956 because it is a worthy example. It is in the CAA dataset but not categorized as COIN by Lyall and Wilson.

Data Sources Reviewed

- Center for Army Analysis "Irregular Warfare" database, 2009
  - Approximately100 "incidences," chosen by availability of data
  - Includes COIN, peacekeeping and other operations
- Lyall-Wilson data set, 2009
  - 286 insurgencies, from 1800-2005; 135 insurgencies in the post-WWII era
  - Cases identified via selection from major academic data sets, e.g., Correlates of War
  - No data on forces actually committed to the COIN operations
- UK MOD (DSTL) "Historical and On-going CT/COIN Campaigns," 2007
  - 44 cases selected via stratified sampling
  - Includes ordinal coding for a broad array of COIN success factors
- Data on US overseas deployments
  - Joint Staff (J-1) in-house database—U.S. forces that deployed for named operations prior to 2000
  - Joint Staff (J-8) and JDS—constructing dataset from Force Management and pay records
- RAND series on Nation Building; Project led by Amb. James Dobbins
  - Separate, detailed reports on U.S., E.U. and U.N. nation-building efforts
- RAND--War by Other Means, Gompert and Gordon, 2009
  - Data base constructed by Martin Libicki; examined 89 insurgencies since World War II

This slide displays all of the 102 cases contained in the CAA dataset, which was the primary data source for this study. The cases selected by IDA for use in our statistical analysis are designated by red bolded type.

As observed on the previous notes slide, Lyall and Wilson identify 57 of the 102 conflicts in the CAA data set as insurgencies. After excluding Western Sahara and those cases for which the CAA data set does not provide sufficient force-sizing data, 41 cases remained for further analysis.

Note that CAA often divided conflicts into separate cases, even though the fighting was continuous. We judged that keeping more than one case from any single conflict would unduly weight them in the statistical analysis.

The CAA Data Set

102 irregular conflicts since 1945. Includes insurgency, terrorism, peacekeeping and other types of operations. **Red Bold text indicates 41 cases used by IDA**.

Aden (Yemen) ‘63-’67
Afghan. (USSR) ‘79-’89
Afghan Civil War I ‘89-’92
Afghan Civil War II ‘92-’96
Afghan Civil War III ‘96-’01
Afghanistan (USA) ‘01-pres.
Algeria I ‘54-’62
Algeria II ‘91-99
Angola ‘61-’74
Angola Civil War ‘75-’88
Angola ‘88-’99
Argentina ‘69-’83
Bolivia ‘67
Borneo (Malaysia) ‘63-’66
Bosnia ‘91-’95
Burundi ‘04-’06
Cabanas (Mexico) ‘67-’74
Cambodia ‘78-’89
Cambodia ‘91-’93
Cameroun ‘55-’60
Chad I ‘65-’69
Chad II ’69-’71
Chad III ’78-’80
Chad IV ‘80-82
Chechnya I ‘94-’96
Chechnya II ‘99-’04
Chechnya III ‘99-pres.

Chile ‘73-’90
Colombia I ‘48-’58
Colombia II ‘58-’64
Colombia III ’64-’09
Congo I ‘60-’64
Congo II ‘00-pres.
Contras (Nicaragua) ‘81-’89
Croatia ‘91-95
Cuba ‘56-59
Cyprus ‘55-’59
Dhofar (Oman) ‘65-’75
El Salvador ‘79-’92
Greece ‘46-’49
Grenada ‘83
Guatemala ‘60-96
Guinea-Bissau ‘63-’74
Hungary ’56
India (Naxalites) ‘67-pres.
Indochina ’45-’54
Indonesia ‘45-’49
Indonesia (E. Timor) ‘75-’99
Iraq ‘03-pres.
Ivory Coast ‘02-pres.
Kashmir ‘88-pres.
Katanga (Congo) ‘61-’63

Kenya ‘54-’60
Kosovo ‘98-’99
Lebanon ‘75-’90
Lebanon ‘90-pres.
Lebanon (Israel) ‘06
Liberia I ‘90-’97
Liberia II ‘03-pres.
Macedonia ‘92-’06
Malaya ‘48-’60
Morocco ‘53-’56
Mozambique I ‘64-’74
Mozambique II ‘75-92
Mozambique ‘92-94
Namibia ‘66-’89
Nepal ‘96-’08
Nicaragua ‘67-’79
Nigeria ‘98-pres.
Northern Ireland ‘68-’98
Oman ‘57-’59
Palestine (UK) ‘44-’49
Palestine(Intifada) ‘87-’93
Palestine ’93-’00
Palestine (Intifada II) ‘00-’05
Palestine ‘05-pres.
Palestine (Gaza) ‘06-08

Panama ‘89
Peru ‘80-’99
Philippines ‘46-’54
Puerto Rico ‘50-’54
Rhodesia I ’66-’72
Rhodesia II ‘72-’79
Rwanda ‘93-’96
Sierra Leone ‘97-’05
Somalia ‘92-93
Somalia ‘06-pres.
Sri Lanka ‘83-’02
Sudan (Darfur) ‘03-pres.
Tibet (China) ‘56-’74
Tunisia ‘52-’56
Turkey ‘84-’99
Uganda (Tanzania) ‘78-’80
Uganda ‘79-’86
Uruguay ‘63-’73
USSR (Ukraine) ’44-54
Venezuela ‘60-’63
Vietnam I ‘57-’60
Vietnam II ‘61-’64
Vietnam III ‘65-’73
Western Sahara ‘73-’91
Yemen ‘62-’70

This slide displays the 41 cases selected by IDA for statistical analysis, showing the force densities and color outcome scorings. (This database is listed in Appendix A.) The solid bars reflect force densities for the AO as determined by our adjustment methodology. The lighter adjacent bars are unadjusted force densities for the entire country. The purpose of including both bars is to demonstrate the effect of applying the adjustment methodology.

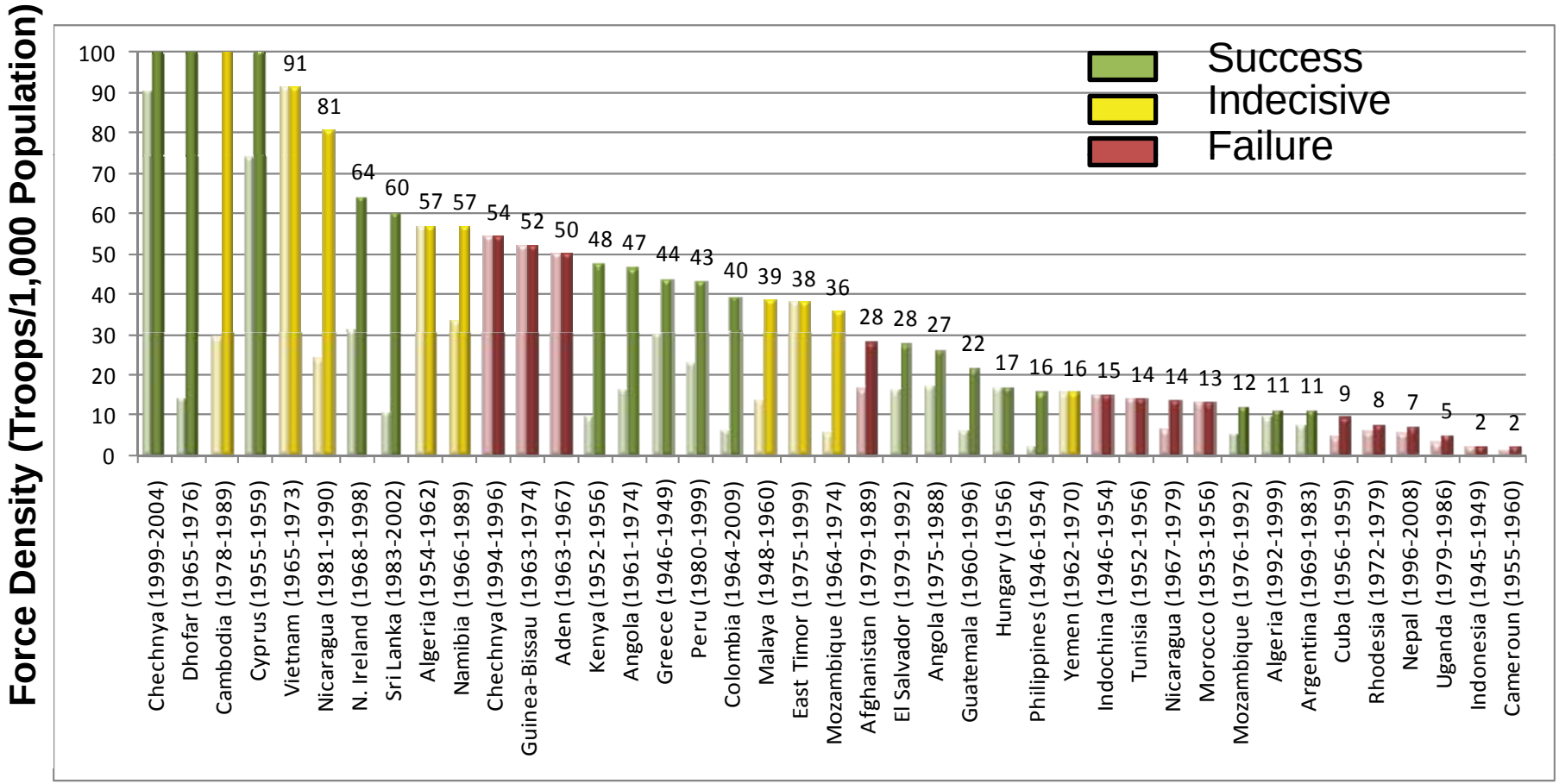
These data indicate a definite trend toward better outcomes for higher (adjusted) force densities, especially if "success" is defined as "no lose." (It is only by using that looser definition of "success" that we were able to obtain statistically significant results, as will be explained in more detail later.) In these "indecisive" cases, the insurgency did not prevail militarily, and in many cases they were clearly defeated. However, the government was unable to convert those operational successes into political "wins." In the case of several colonial wars, the colonial power made a decision to withdraw even though the insurgency was largely defeated – Algeria (1954-1962), Angola (1961-1974), and Mozambique (1964-1974). (The outcome scoring is of prime importance to our statistical analysis and is explicated further in Appendix A).

Do such historical conflicts provide a valid basis for estimating requirements for future conflicts involving U.S. forces? Although every conflict is unique, there are many aspects that are common. General David Petraeus, makes the point, writing in the foreword to FM 3-24: "You cannot fight former Saddamists or Islamic extremists the same way [we] fought the Viet Cong…Nonetheless, all insurgencies, even today's highly adaptable strains, remain wars amongst the people. They use variations of standard themes and adhere to elements of a recognizable revolutionary campaign plan." This is the basic reason that we consider the history of insurgency to be a plausible basis for generalizing about its future.

Nonetheless, any such statistical analysis rests on the assumption that historical cases are sufficiently comparable to render the observed correlations meaningful. But how can that assumption be confirmed? Should the British campaign to win hearts and minds in Malaya be part of the same data set as the Russians' brutal assault on the people of Chechnya? Questions such as these represent an important limitation of statistical analysis, especially when the size of the dataset is limited.

Force Densities for Reduced Areas of Actual Stability Operations—IDA-selected subset of 41 from CAA Data

- Mean force density is 57 troops/K; range is 655-2 (Standard Deviation = 107)
- Removal of top two outliers brings mean down to 36 (Standard Deviation = 30)
- For force densities >20, 80% were wins or "indecisive"
- For force densities <16, 77% were losses



This chart summarizes graphically the data given in detail on the previous chart. The graph on the left shows the observed fractions of “win” and “no lose” versus force densities in the area of operations. The graph on the right shows the impact of using force densities for the AO versus the entire country for the “no lose” outcome.

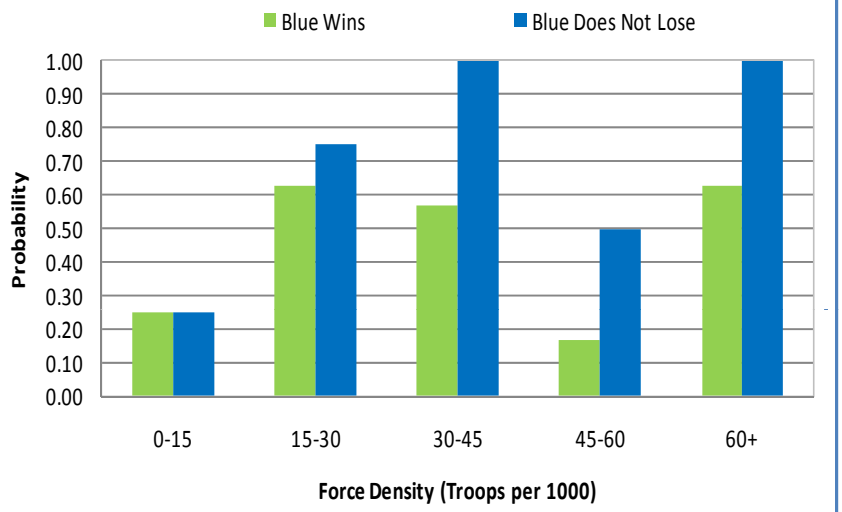
Using reduced populations, simply taking a force density (FD) of 20 as a point of division, the 25 cases with FDs greater than 20 show four losses (16%), while the 16 cases with FDs below 20 show 10 losses (62.5%). For the 14 cases with FD less than 15, 71% were losses. These simple statistics themselves would seem to establish a case for a relationship between force density (in the area of operations) and better outcomes. Obviously, the three losses all with FDs in the 50-60 range illustrate that force density alone cannot completely explain the outcomes (see Appendix A). This should be no surprise whatever to anyone well-informed about this type of conflict. As stressed earlier, many other factors, not the least of which are strategy and tactics and political factors, are frequently dominant.

Working independently, three sets of authors have conducted sophisticated analyses of insurgency/COIN outcomes over the past several years. The analyses are not directly comparable, since the authors each assembled their own data sets. However, all three studies rely on the same statistical approach – *logistic regression* – to search for meaningful patterns in the data.

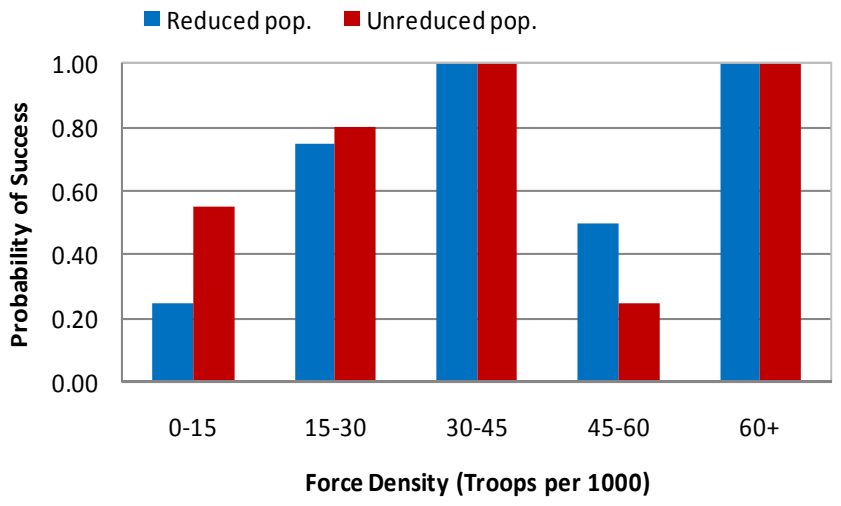
Logistic regression (the mathematical formulation is seen in the chart following) is the preferred technique in numerous disciplines for studying binary outcomes, *i.e.* outcomes that are fully captured by answering a yes-or-no question such as “Did this individual suffer a heart attack?” or “Did this bridge collapse?” The logistic regression approach accommodates consideration of both discrete and continuous independent variables as candidate explanatory variables for the observed outcomes. For example, a study of heart attacks can assess the influence of both one’s age (a continuous variable) as well as whether or not one smokes (a binary variable). If one collected data about a set of patients, indicating whether they had heart attacks, whether they smoked, and how old they were, logistic regression analysis could determine whether there is a statistically significant relationship between the candidate variables (age, smoking) and the outcome (heart attacks). It could also generate an equation expressing a relationship between age and smoking habits (independent variables) and the probability of having a heart attack (dependent variable).

Force Density vs. Observed Rate of Success in IDA-Selected Data

Comparing "win" vs. "no lose" with Population Reductions



Comparing "no lose" with and without Population Reductions



We chose to use "no lose" to define "success" in our statistical analysis

- Criteria for defining "win" vs. "no lose" are loose—categorizations differ significantly in the several datasets examined
- "No Lose" includes both victories and indecisive outcomes
  - Indecisive outcomes generally involve significant operational success for COIN forces, but without a clear cut strategic or political success
  - Our interest is in what force density is required to achieve operational success; strategic success is likely to be determined more by political factors
- Using "no lose" versus "win" yielded statistically significant results, otherwise it failed
- Without adjusting force densities to the actual AO, statistical significance also failed

E. Statistical Analysis Using Logistic Regression

This chart shows the basic mathematical construct for the logistic regression technique that was used for our statistical analysis of the significance of several variables in predicting outcomes for stability/COIN operations. (For an introduction to logistic regression, see David W. Hosmer and Stanley Lemeshow, *Applied Logistic Regression*,2nd edition, New Jersey: John Wiley & Sons, 2000.) This technique, which differs mathematically from the more familiar linear regression techniques, involves the iterative solution of non-linear equations. It enables a determination of statistical relationships between one or more continuous or finite-valued independent variables to the probability of an outcome.

Is logistic regression appropriate for studying insurgency/COIN outcomes? The answer depends on whether the outcome of an insurgency can be clearly described as a victory for one side and a defeat for the other. The studies reviewed here all agree that insurgency/COIN campaigns have three possible outcomes: success, failure, or something in-between (which we call indecisive). The in-between category includes stalemates, negotiated outcomes and other indecisive results. One way to address this problem is to exclude all indecisive outcomes from the regression. However, when working with a small set of cases, this may result in the loss of too much data. Another option is to run the regression twice, first counting indecisive outcomes as a success for the insurgents, then counting them as a success for the COIN force. It is also possible to use a more general logistics regression technique that permits an ordered set of more than two outcomes. That approach has been used by Andrew Hossack of the United Kingdom Defence Science and Technology Lab (DSTL). That approach would be preferred if sufficient data were available to permit a finer discrimination of outcome.

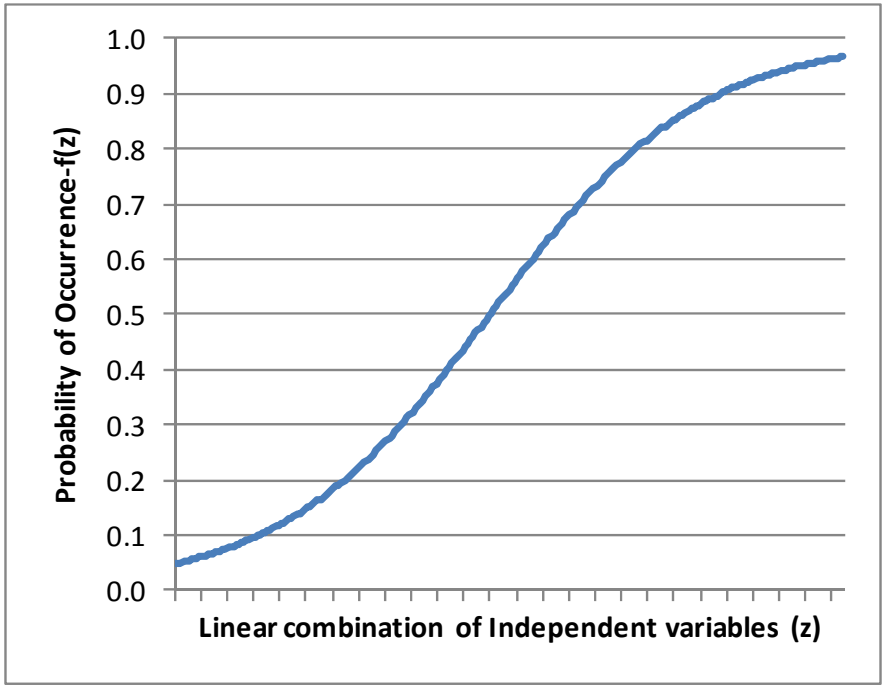
Logistic Regression

- Used to predict the probability of occurrence of an event by fitting data to a logistic curve; it is a generalized linear model used for binomial regression

$$f(z) = \frac{1}{1 + e^{-z}}$$

$$z = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \cdots + \beta_k x_k,$$

The x_i *s* are the independent variables. f*(z)* is the probability of occurrence, given *z* (determined by given x_i *s*) . The regression model estimates the betas.



This chart shows the basic result of applying the logistic regression technique. Force density was found to be statistically significant (more discussion later) and the relationship between force density and probability of success (defined as Blue not losing) is seen in the graph.

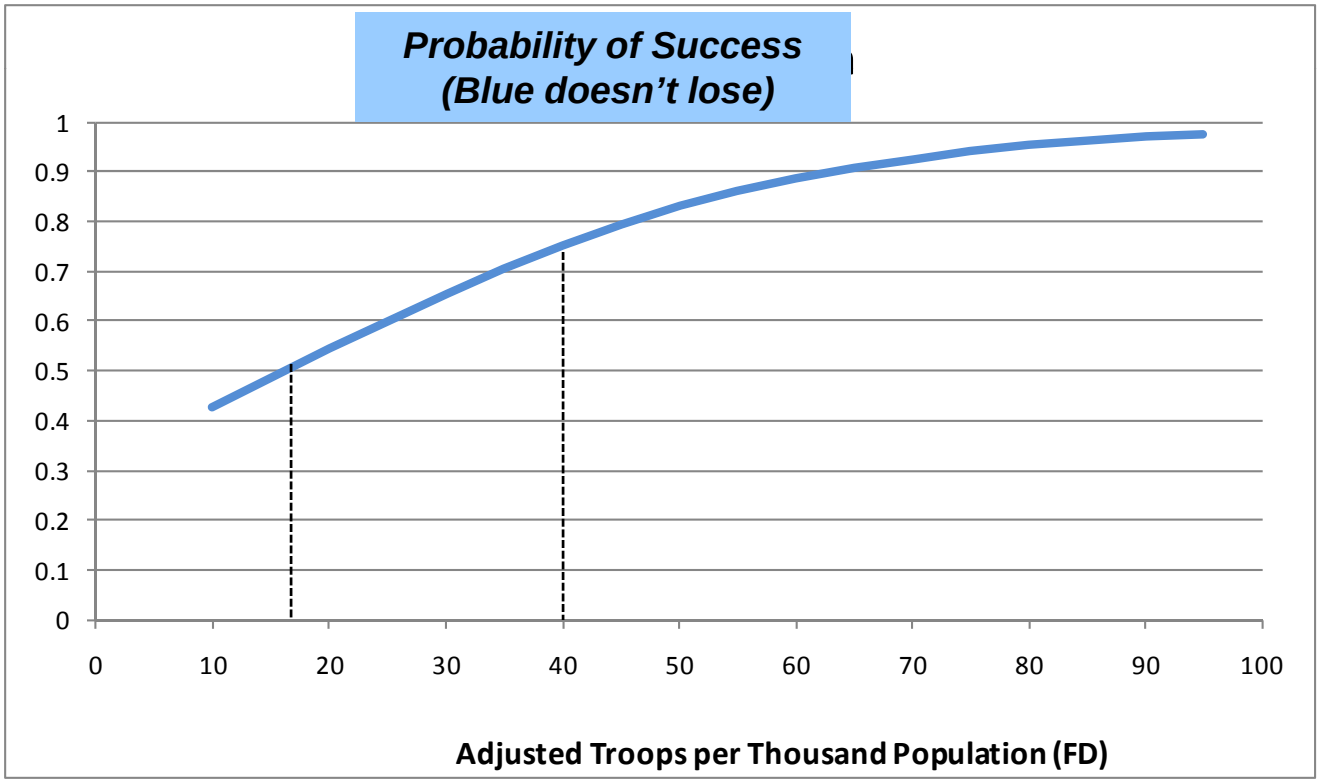
This result confirms FM 3-24's citation of a force density of 20 per 1,000 population in the AO as that "often considered minimum troop density required for effective COIN operations" in that such a force density corresponds to a 54% probability of success.

Our data further suggest that one way to improve the odds of success in COIN operations is with a higher force density. For the 18 conflicts with force densities 40 or greater, 15 were deemed successful (83%). However, for the eight cases with force densities in the 20-40 range, only four (50%) were deemed successful. Thus the suggestion (implied) in FM 3-24 that 20-25 troops per 1,000 is *an appropriate range of force densities* for COIN operations was not confirmed by the study. A more appropriate range, based on these data, might be 20-50 troops per thousand—an FD of 50 corresponds to an 80% probability of success using the logistics regression.

An extremely important caveat in using this approach is the implicit assumption that all conflicts are, in some sense, equal except for force density. If a COIN force is unable to achieve a force density of 40-50 troops per 1,000 inhabitants, it may still improve its odds of success by employing superior strategy and tactics. Non-military factors, such as holding elections or reducing unemployment, can also affect the odds of success. The data currently available did not allow us to capture such relationships. In fact, some important relationships may not be quantifiable at all. As FM 3-24 states , everything may remain "very dependent upon the situation."

Logistic Regression Predictive Curve: Force Density versus Probability of Success

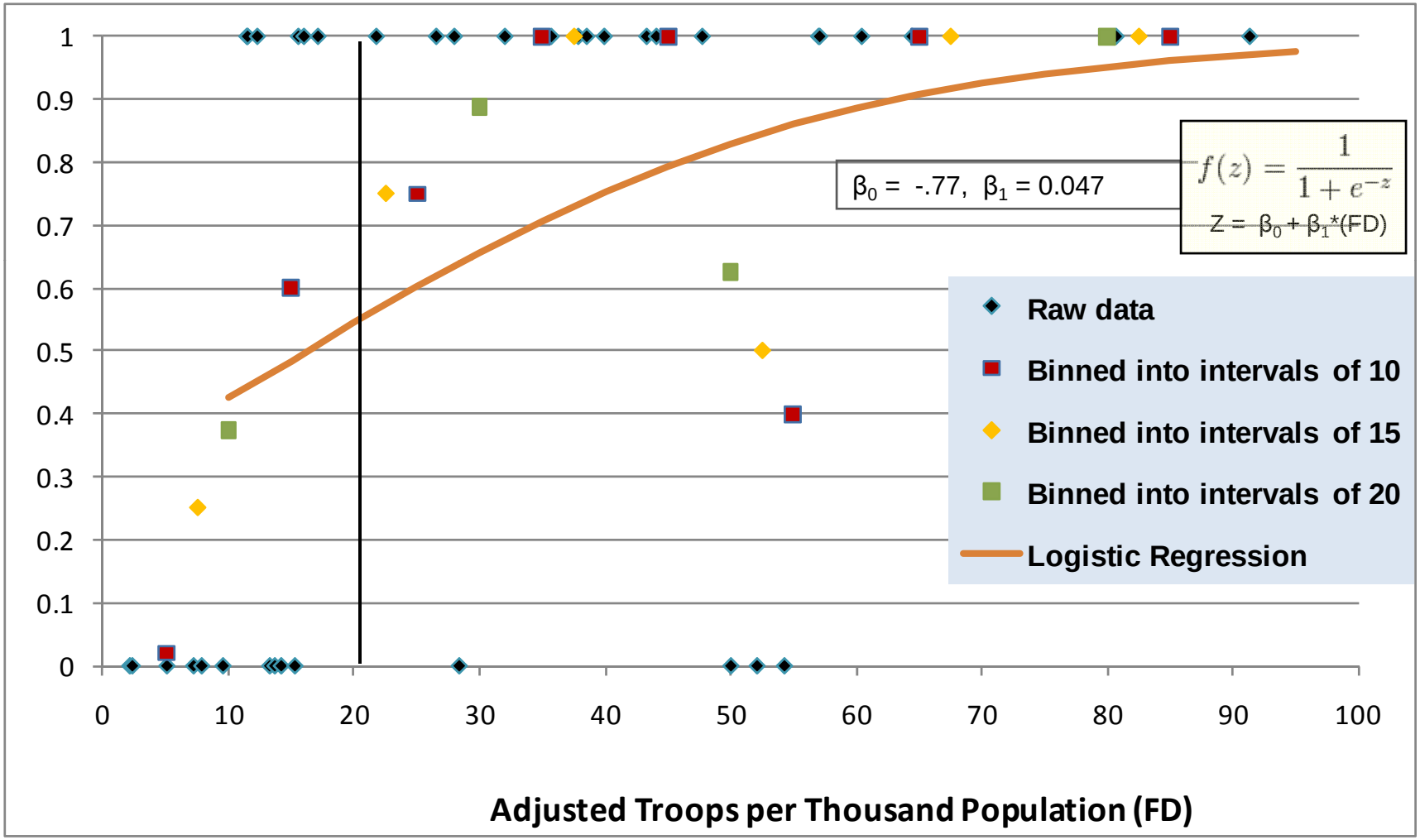
- IDA-selected subset of CAA data
- Regression results indicate probability of success (Blue doesn't lose) of about 50% for a COIN force with a force density 16 troops per thousand
- Probability of success rises above 75% with force densities greater than 40 troops per thousand



This chart shows the logistic regression displayed with the actual data points and with the data binned into various intervals (the process used by the logistic regression model). This display is useful for understanding the limitations of the statistical analysis. The finer intervals provide more data points to fit to, but because the number of events in each interval is progressively smaller, the observed probabilities have greater variability. The coarser intervals, on the other hand, provide more reliable probabilities but fewer data points. The logistic curve appears to be a reasonable fit using intervals of 20; however there are only four data points. Looking at the intervals of 10 or 15 provides more data points, but the sparseness of the data within the cells reduces the reliability of the observed probabilities.

The regression coefficients ($\beta_0 = -0.766$, $\beta_1 = 0.047$) are in reference to the regression equation shown on page 51.

Logistic Regression Predictive Curve with Raw Data and Binnings



This chart summarizes the results of our statistical analysis. As already noted, we found that the relationship between force density and conflict outcome was statistically significant. However, in several ways that finding is not particularly robust. It depends on three choices that we made: The first was to use force densities adjusted for the populations corresponding to our estimates of the actual AO (rather than the entire country) in historical campaigns. The second was to count indecisive outcomes as military operational successes for Blue. The third was the rescoring of several important outcomes as indecisive, rather than failure for Blue. If *any* of those choices are not made, the relationship does not prove to be statistically significant.

The first choice—to use populations reduced to the actual AO—seems incontrovertible—the only reservation being whether the actual AO can be accurately identified from the available information. The second and third choices, however, are more debatable, so that topic is explored in greater detail in Appendix A.

Our results differ from the findings of other studies. In fact, we found no other study of insurgency/COIN outcomes that identified a relationship between force density and outcomes. Only two other studies performed statistical analysis and neither found a statistically significant relationship between force density and outcome. Both studies, however, found a statistically significant relationship between COIN-to-insurgent force ratio and outcomes. Hossack found that COIN forces can double their odds of success by increasing the ratio of COIN forces to insurgents by a factor of four. Justine Blaho and Lisa Kaiser of CAA found that increasing the ratio of troops to insurgents has a much smaller effect, although still a significant one. Blaho and Kaiser observe that there is “little accuracy” to the information available about how many insurgents participated in any particular conflict. Our concerns about this inaccuracy are even greater. We do not believe that information about the size of insurgent forces is sufficiently credible to be subjected to statistical analysis. (See Appendix C for full references for the cited studies.)

And even if it were, it is not a satisfactory explanatory variable for our purposes because it is not possible to estimate with any reliability the size of a postulated insurgency force for a hypothetical conflict in a particular country many years into the future. Of course scenarios could be postulated that could include that information. However, our objective is to develop force size estimates for a wide range of countries for which no such scenarios have been developed.

Statistical Analysis of CAA Data

- Analysis used logistic regression performed on subset of 41 cases, including COIN *only*
- Force density is not significant when calculated using raw figures for national populations
- Force density *is clearly* significant when calculated using force densities for the *actual areas of operation*
  - Force density is significant for both single regression and for multiple regression including other leading variables
  - Caveat: Significance only obtains when the dependent variable is "Blue does not lose", i.e. indecisive outcomes are counted as a success for Blue
- Variables found by CAA to be statistically significant were not found to be so by this study, using our reduced dataset, adjusted force densities, and IDA outcome scoring

This chart displays the *p* values for force density and four other variables identified by CAA as having a significant impact on COIN outcomes. In effect, a *p* value measures the likelihood that an apparent correlation between variables is random, rather than significant. Statisticians generally consider 5% (.05) to be the threshold for significance. A *p* value *less than* .05 qualifies a statistical relationship as significant. (However, some analysts, especially in the social sciences, use a looser criterion of 10%, meaning that the user of such results is willing to accept a 10% that the observed relationship is due purely to chance. Both CAA and DSTL analysts used a 10% criterion)

The *p* values are from multivariate regressions, meaning that the impact of all five variables is measured simultaneously. It is also possible to run single-variable regressions that test for the impact of only one variable at a time. This may provide an initial indication of whether a variable is significant, whereas multiple regression shows whether a variable that is significant by itself is overpowered when other variable are considered.

The table in the chart displays the outcomes for two multiple regressions with different dependent variables: "Blue Wins" and "Blue Does Not Lose." As noted before, force density (adjusted for the actual AO) correlates significantly with Blue avoiding loss, but not with Blue achieving a win.

One other variable was found to have a statistically significant relationship at the 5% level—whether the Blue force included significant forces from a developed nation. In that case, the coefficient is negative, indicating that intervention by a Western force reduces the chance of success. The history of COIN operations suggests that several factors may contribute to this effect. First, there is a tendency for developed nations to deploy combat forces only after other forms of assistance have failed. And, Western forces often have to overcome a cultural barrier to effective operations, and may be perceived as an occupying or colonial power, thus decreasing their legitimacy (an effect observed in both Iraq and Afghanistan). Statistically, the correlation between failure and the presence of Western forces indicates that Western forces must achieve a higher force density to enjoy the same probability of success as an (equally capable) exclusively indigenous force. Whereas the relationship between force density and improved outcomes is fairly straightforward, the impact of the introduction of Western forces is more subtle and problematic.

Statistical Analysis of IDA-Selected Subset of CAA Data with Adjusted Force Densities

- Table displays p-values for logistic multiple regression with ***key variables*** identified by CAA or IDA
- Variables with a p-value of ***less than .05*** are considered statistically significant

Variable	Blue Wins	Blue Does Not Lose
Force Density (Adjusted)	.251	**.011**
Force Ratio	.080	.311
Number Red Factions	.977	.833
Motivation = Nationalism	.091	.758
Blue = Developed Nation	.126	**.028**

An earlier generation of analysts emphasized force ratio as the appropriate metric for determining the number of boots on the ground. For example, David Galula asserted that the French failure in Indochina was inevitable because the counterinsurgents never achieved a force ratio of at least ten to one. (David Galula, *Counterinsurgency Warfare: Theory and Practice* (1964 [2005]), 32)

Galula's position is especially interesting because he is best known as an advocate of population-centric COIN. In contrast, today's advocates of population-centric approaches (such as the authors of FM 3-24) consider force density the relevant metric, because it prioritizes the relationship of the military to the population, rather than the relationship of the military to the insurgents. Since the publication of James Quinlivan's seminal article on force density in *Parameters* more than a decade ago, force density has stood alone as the accepted metric for gauging the sufficiency of COIN forces. (James T. Quinlivan, "Force Requirements in Stability Operations," *Parameters* 24:4, 59-69.)

Given this background, the CAA and DSTL findings with regard to the statistical significance of force ratios represent a challenge to the conventional wisdom. A recent study by RAND also identified force ratio as a major influence on COIN outcomes, although that study did not employ any measures of statistical significance. (David Gompert & John Gordon IV, *War By Other Means: Building Complete and Balanced Capabilities for Counterinsurgency* (RAND Corporation, 2008), 373-395.)

Although we do not take issue with the analytical techniques employ by CAA, DSTL and RAND, we do question the validity of the underlying data. Historical data on insurgents' numerical strength consists mainly of "guesstimates" by their opponents. The U.S. military's experience in Vietnam and Iraq suggests that such estimates may not rest on substantial evidence and may be subject to large margins of error. Moreover, there is still no consensus regarding who should even be counted as an insurgent. Only full-time professionals? The teenager paid to bury one or two improvised explosive devices? Women who smuggle supplies? In light of these uncertainties, we consider force ratio to be a flawed and unreliable indicator.

Findings on Force Ratio

- Whereas force density is the ratio of troops to population, force ratio is the number of counterinsurgents per insurgent
- Previously considered by experts an essential measure of force adequacy, but not supported by FM 3-24
- CAA and DSTL both found statistically significant relationships between force ratio and success in COIN operations
- This study's statistical analysis shows force ratio being on the threshold of statistical significance for Blue Wins. (See previous slide)
- Insurgent strength estimates in COIN campaigns are notoriously unreliable
  - Good intelligence on insurgent strength is difficult to obtain
  - "Membership" in an insurgency is extremely fluid
  - In half the cases, CAA and DSTL figures for insurgent manpower differ by a factor of two

This chart summarizes the conclusions from our statistical analysis.

Conclusions from Statistical Analysis of Historical Data

- Statistical analysis of 41 selected historical incidences using *logistic regression* yielded important insights into the significance of *force densities* to outcomes in stability operations
  - The data indicate that the fraction of successful outcomes decreases greatly for force densities below 15 troops per thousand
  - Using force densities for the *actual* areas of operation, statistical analysis provides evidence that:
    - The rough rule of thumb of 20-25 troops per thousand population is consistent with about a 50-60% chance of success
    - Troop density is significant in correlating (positively) with successful COIN outcomes
    - For operations with ratios above *40*, the statistically predicted success rate exceeded 75%

The analysis supports the 20 troops per thousand rule of thumb as a threshold for success but not 20-25 as an appropriate range to consider

F. Comparison to Findings from Other Research Organizations

When our preliminary results were presented to our sponsor, there was concern that our findings differed from those of other organizations that have conducted similar research. Thus the sponsor requested that IDA host a seminar with the objective of better understanding those differences.

On December 14, 2009, IDA hosted a seminar for researchers from five analytical organizations to review the results of their recent and ongoing research in the area of force sizing for stability and COIN operations. First, there was a major difference with regard to whether force density or force ratio is the appropriate metric for sizing a COIN force. This disagreement is the latest round of a debate that began more than forty years ago, when COIN operations first became the subject of rigorous study.

The focus of the next two slides will be on the sources of uncertainty in data and methodology that are responsible for the divergent results of recent research. The first major source of uncertainty is the data sample employed by different researchers. Four of the projects discussed at the seminar entailed cross-national studies of from 40 to 90 conflicts that have taken place since World War II. Although none of the studies looked at precisely the same set of conflicts, there is substantial overlap in the selection of cases. Second, there was significant disagreement about the scoring of the outcome of the conflicts. Third, the different studies drew on separate estimates of how many counterinsurgent and (especially) insurgent forces participated in those conflicts. Finally, the four studies took different approaches to demarcating the boundaries of the AO in which the conflict took place. Of the four factors, the one that contributed most to divergent research results was the lack of consensus on how to score conflict outcomes.

Comparison to Findings from Other Research Organizations

- Our findings differ significantly from those of similar studies by CAA, DSTL, and RAND
- Those studies all concluded that force ratio, but not force density, influenced COIN campaign outcomes
- Differences in data and methodology responsible for these conflicting results include:
  - Selection of data sample--focus on COIN or on all stability operations?
  - Scoring of campaign outcomes
  - Historical forces data--Blue forces are an area of relative agreement; Red force strength is much more uncertain
  - Defining the AO—only done by one other study (DSTL) and differences could be resolved

Lack of consensus on how to score conflict outcomes emerged as the primary explanation for differences in findings

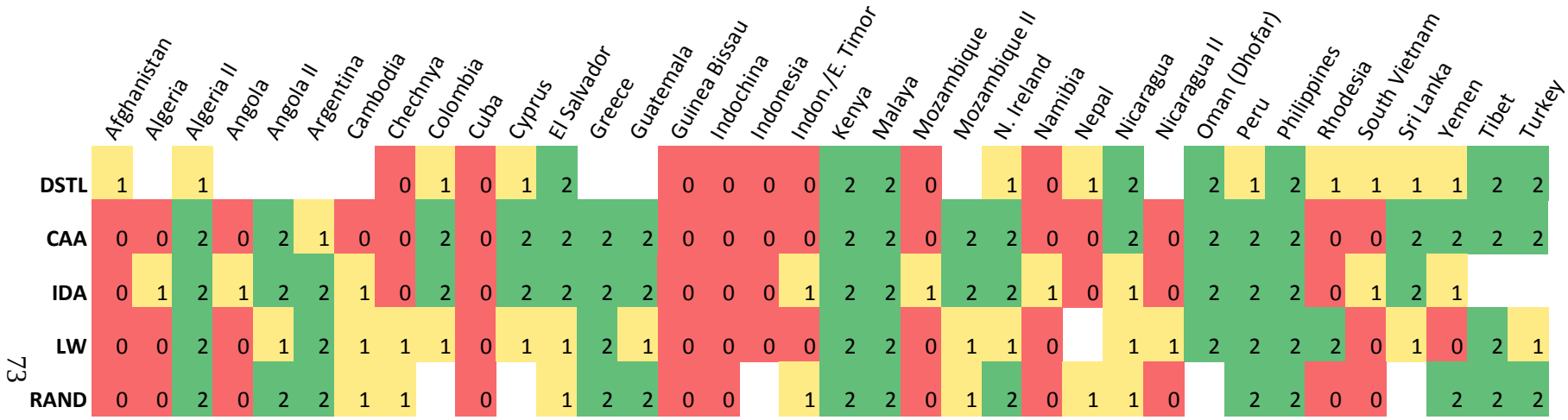
What counts as success for COIN operations? There is a firm consensus that the outcome of many operations cannot be described as either victory or defeat. Analysts label these intermediate outcomes as "mixed," "indecisive," "draws" or the like. What analysts have not succeeded in delineating is whether meaningful thresholds exist that distinguish decisive outcomes – successes and failures – from the indecisive outcomes. Jason Lyall and Isaiah Wilson write that "a draw occurs when [a government] is forced to concede some, but not all, insurgent demands, and neither side obtains its maximal aims." How much is 'some'? If a government offers amnesty to insurgent fighters, is that a significant concession, or a means of consolidating its victory?

The lack of clear answers to such questions has resulted in analytical confusion. A comparison of the scorings assigned by this study and the others demonstrates that there is significant disagreement about the threshold between decisive and indecisive outcomes in actual historical cases. This lack of agreement is problematic, because the validity of logistic regression depends on the presence of a clear-cut outcome as the dependent variable. When working with small data sets, re-scoring the outcome for just a few cases can reverse the results of statistical analysis. For example, our finding that force density has a significant impact on campaign outcomes depends on the scoring of five cases as indecisive, even though many others code them as failures.

Disagreement about campaign outcomes can be difficult to resolve. Historians continue to debate whether the U.S. lost militarily in Vietnam, or suffered a political defeat. Similar debates persist about the outcome of other conflicts, such as the French in Algeria. Thus, there will always be intrinsic limits to the value of studies that rely on simple "win, lose or draw" assessments of COIN outcomes. (See Appendix A for a discussion of our scoring of outcomes for the conflicts in our dataset.)

Given how few campaigns are available for inclusion in COIN data sets, disagreements about a small number of cases can have decisive impact on statistical analyses. This argues for expanding the dataset (if possible) and/or better understanding the outcomes and the drivers thereof.

Differences in Scoring Campaign Outcomes



- 36 conflict cases scored by at least 4 out of 5 organizations
  - Complete agreement on outcome in only 11 cases
  - Single dissent in an additional 12 cases
  - Divided opinion in remaining 13 cases
  - No consensus on how much concession is allowed before one side can no longer be called the outright winner

- *There is no standard for what should count as a win, loss, or indecisive*
- *With small sample sizes, re-scoring several cases can change the outcome of statistical analysis*

G. Linear Regression Analysis

Given that we have established a statistically significant relationship between force density and COIN conflict outcomes, we can use linear regression techniques for further insights into the size of a force for a country for which a stability operation may be required in the future. The linear regression technique is applied to the set of 27 contingencies in our sample in which the "Blue" force did not lose, plus three more recent campaigns (Bosnia, Kosovo, and Iraq) that are also characterized as such. The linear regressions consider population as the independent variable and force density as the dependent variable. In effect, we are asking how the force density for "successful" COIN operations has varied historically with the number of inhabitants in the AO.

This approach is not as statistically rigorous as logistic regression. It ignores the failures, and there is an implicit assumption that the force achieving success was correctly sized. Despite its drawbacks, the method still offers useful insights, and has the advantage of providing force projections that seem more "reasonable," in that they are closer to results from the MFSM and more consistent with recent experience in Iraq. In light of those considerations, we chose this method as a baseline in *illustrating* results.

This chart displays the "successful" cases selected for the linear regression analysis. The display shows both the reduced and unreduced populations and corresponding force densities; however, only the reduced values were used for the linear regressions.

The results show a statistically significant relationship between force density and the population of the AO. The *p* value for this relationship is less than 1%.

Despite having an extremely low *p* value, the data here have a high standard deviation (greater than the mean), indicating that the force density for a specific campaign may depart substantially from the mean. Nonetheless, it does mean that estimating the required force density for a *future* operation will be subject to considerable uncertainty. As FM 3-24 emphasizes, conditions on the ground will likely necessitate major adjustments.

Dataset for Linear Regressions

Thirty conflicts selected as the baseline for projections

- All 27 cases from IDA subset of the CAA dataset with "win" or "no lose" outcomes
- Added Iraq (OIF), Bosnia, and Kosovo
- Removed one extreme outlier (Western Sahara, 1973-1991, force density of 655 troops/K)

Conflict (Years)	Popula-tion (M)	Troop Density	Adjusted popula-tion	Adjusted Troop Density	Conflict (Years)	Popula-tion (M)	Troop Density	Adjusted popula-tion	Adjusted Troop Density
Greece (1946-1949)	7.6	30.5	5.2	44.1	Cambodia (1978-1989)	8.7	29.8	1.9	138.6
Philippines (1946-1954)	21.5	2.6	3.6	16.0	El Salvador (1979-1992)	5.0	16.4	3.0	27.9
Malaya (1948-1960)	5.5	19.2	2.8	38.5	Angola (1975-1988)	10.0	17.5	6.6	26.5
Kenya (1952-1956)	7.0	10.2	1.5	47.7	Nicaragua (1981-1990)	3.7	24.6	1.1	80.7
Hungary (1956)	9.9	17.1	9.9	17.1	Mozambique (1976-1992)	14.2	5.8	6.7	12.2
Cyprus (1955-1959)	0.6	74.6	0.4	100.9	Guatemala (1960-1996)	9.7	7.0	3.1	21.7
Algeria (1954-1962)	11.0	57.1	11.0	57.1	Peru (1980-1999)	23.5	6.4	12.7	11.8
Yemen (1962-1970)	4.0	15.5	4.0	15.5	Chechnya (1995-2009)	13.7	7.3	0.9	116.0
Vietnam (1965-1973)	17.9	62.7	17.9	62.7	Sri Lanka (1983-2002)	18.9	10.9	3.4	60.4
N. Ireland (1968-1998)	1.5	31.5	0.8	64.5	East Timor (1975-1999)	0.8	37.8	0.8	37.8
Mozambique (1964-1974)	8.8	5.9	1.4	35.7	Algeria (1992-1999)	30.1	10.1	26.5	11.5
Angola (1961-1974)	6.5	10.5	2.1	32.0	Colombia (1964-2009)	45.6	6.7	7.7	39.9
Dhofar (1965-1976)	0.9	14.7	0.1	270.0	Bosnia*	3.4	16.8	3.4	16.8
Argentina (1969-1983)	28.5	8.0	20.0	11.4	Kosovo*	2.0	23.7	2.0	23.7
Namibia (1966-1989)	1.1	33.2	0.7	57.0	Iraq*	26.0	25.4	23.0	28.7

** Not included in the logistic regression dataset*

Green indicates "win"
Yelloe indicates "indecisive"

Average	50.8
Median	36.8
Standard Deviation	52.2

This chart displays the results of the linear regression analysis. As anticipated, the fit is less than ideal, because individual case force densities often lie at a considerable distance from the regression line. However, the correlation coefficient (R^2) for our equations is roughly 0.45, indicating that the equation accounts for 45% of the variation. This is a relatively good score for complex social enterprises such as COIN operations.

It should be noted that we ran a separate regression for countries with populations less than two million. The main reason for doing so was to reduce the distortion caused by the constant term. If we had used the regression on the full set for smaller conflicts, the 48,500 constant would have dominated the estimates and yielded unreasonably large results for smaller countries. (The effect on our final results was very limited since we are primarily interested in the force requirements for larger nations.)

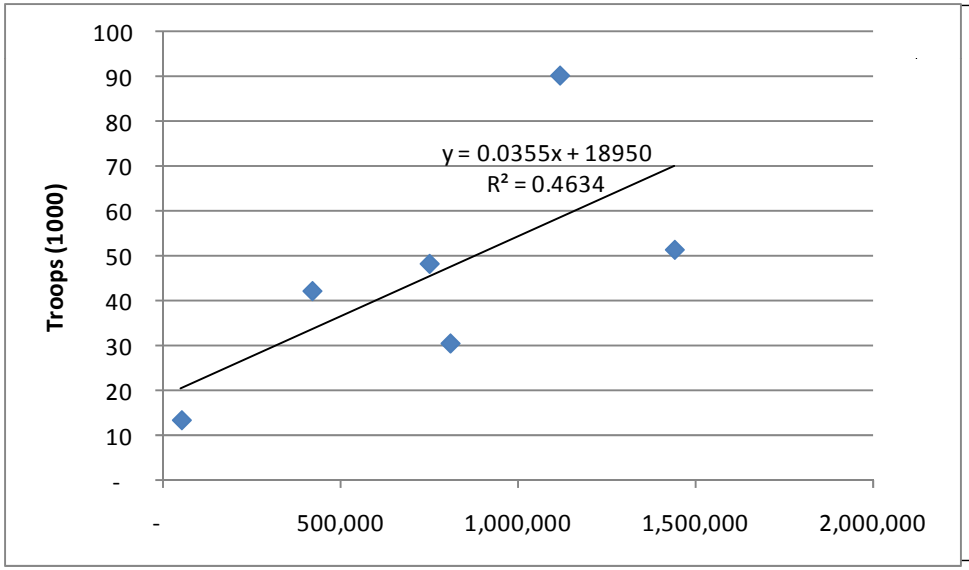
It is also possible to take the results of the linear regression for a given country and ask what probability of success that force density corresponds to in the logistic regression curve. (Because of the constant terms the force densities vary depending on the population of the AO for the country.) The results vary from a low of 54% for the largest country to 84% for the smallest, averaging 61% with a median of 59%.

Linear Regression Equations

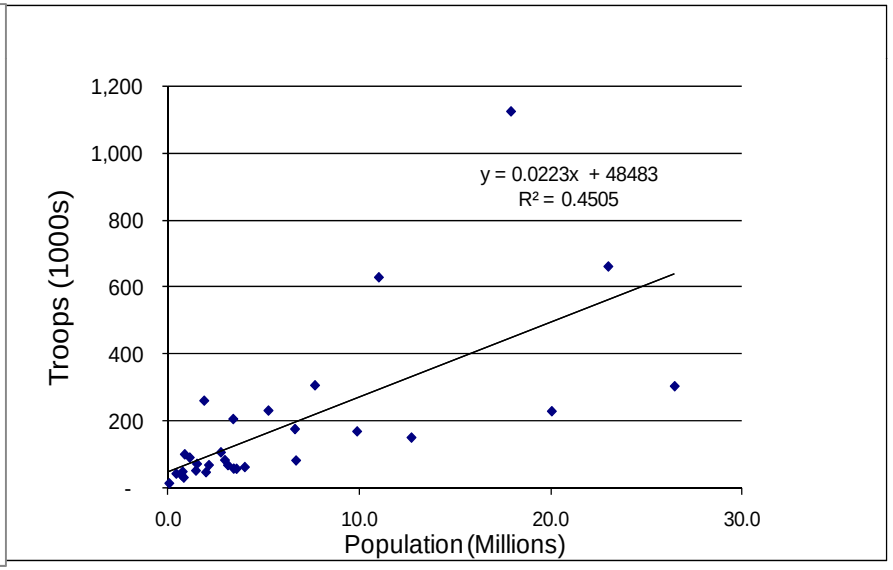
Populations < 2.2 million: Troop Strength = .0355*(population) + 18,950

Populations> 2.2 million: Troop Strength = .022*(population) + 48,480

Small Wars



All Wars



Equation for entire set was used (not just population >2.1M)

This analysis provides an estimate of 22 troops per thousand for larger countries, plus a population-independent constant factor

H. Comparison of Force Size Projection Methodologies

This chart compares the results using the three methodologies considered averaged over the set of 54 countries for which Blue force estimates were made. The purpose is merely to illustrate the impact of using the different methodologies, since there is little analytical meaning to these average *per se*. There are four entries on the chart because we use our logistic regression methodology to project the Blue force requirement for both a 50% and 75% probability of success.

The left axis is the estimate of the total Blue force size, while the right axis displays the force density in troops per 1,000.

As the chart shows, the lowest projection using our new methodology provides an estimate about 30% higher than the updated MFSM.

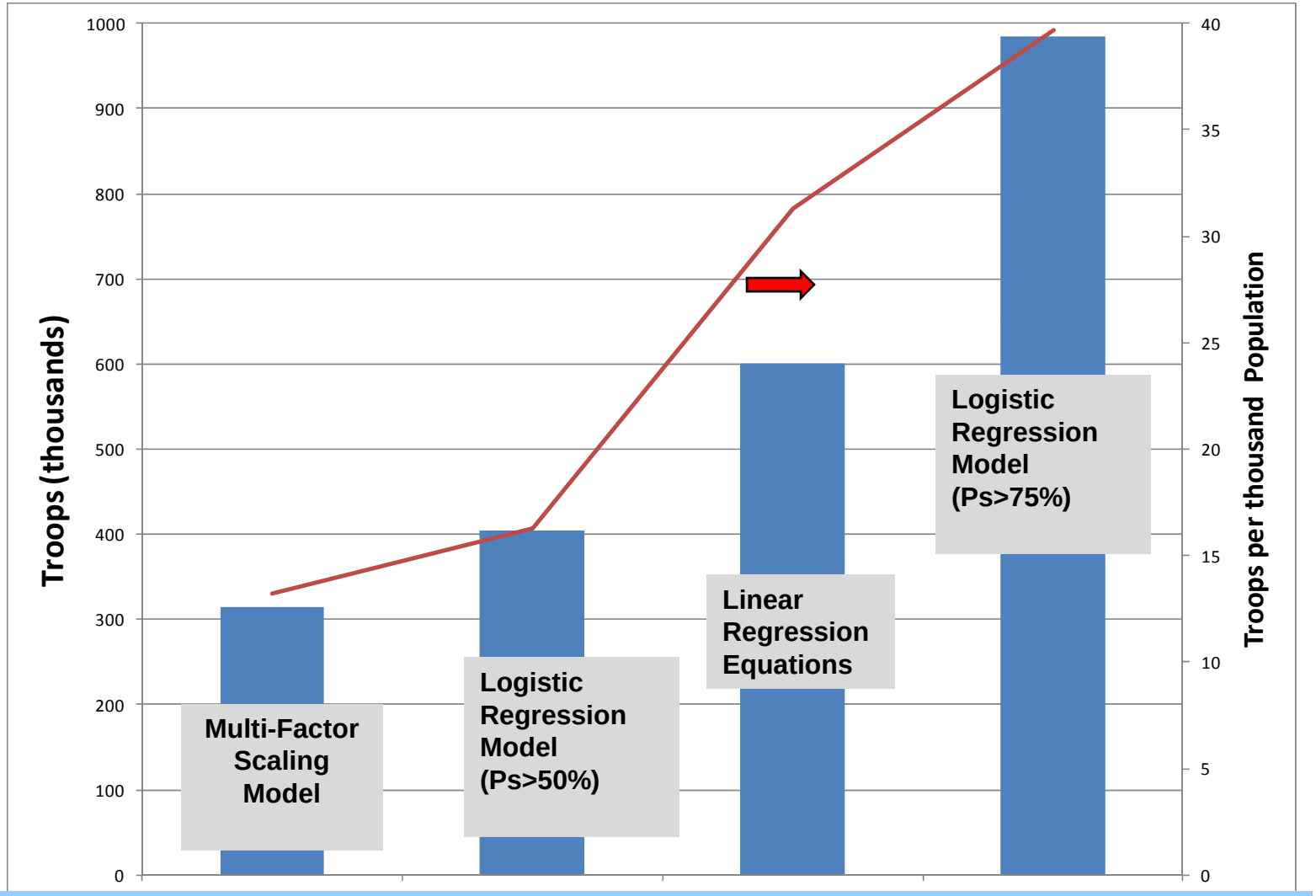
For the logistic regression methodology, the projected force density will remain constant regardless of which countries are used as illustrations. This is true because population is the only parameter that determines the value of the projections. A 50% chance of success will always require about16 troops per 1,000 inhabitants in the AO. A 75% chance of success requires about 40 troops per 1,000.

For the linear regression methodology, the projected force density (for larger countries) will always be 22.3 troops per 1,000 inhabitants plus the value of the constant divided by the population of the country (or countries) of interest. For an AO with 3 million inhabitants, the force density would be 38 troops per 1,000 inhabitants. For an AO with 30 million inhabitants, the force density would be 24 troops per 1,000. For larger countries, the probability of success based on the logistic regression is about 57%, increasing to 80% for smaller countries. (The average is 66%)

For the projections generated by the MFSM, the effect of scaling on population, while still playing an important role, is tempered by the other variables included, as well as by the values for the reference operation used (Iraq—OIF).

Comparison of Force Size Projection Methodologies

Average Force Size and Troops per Thousand for 54 Countries



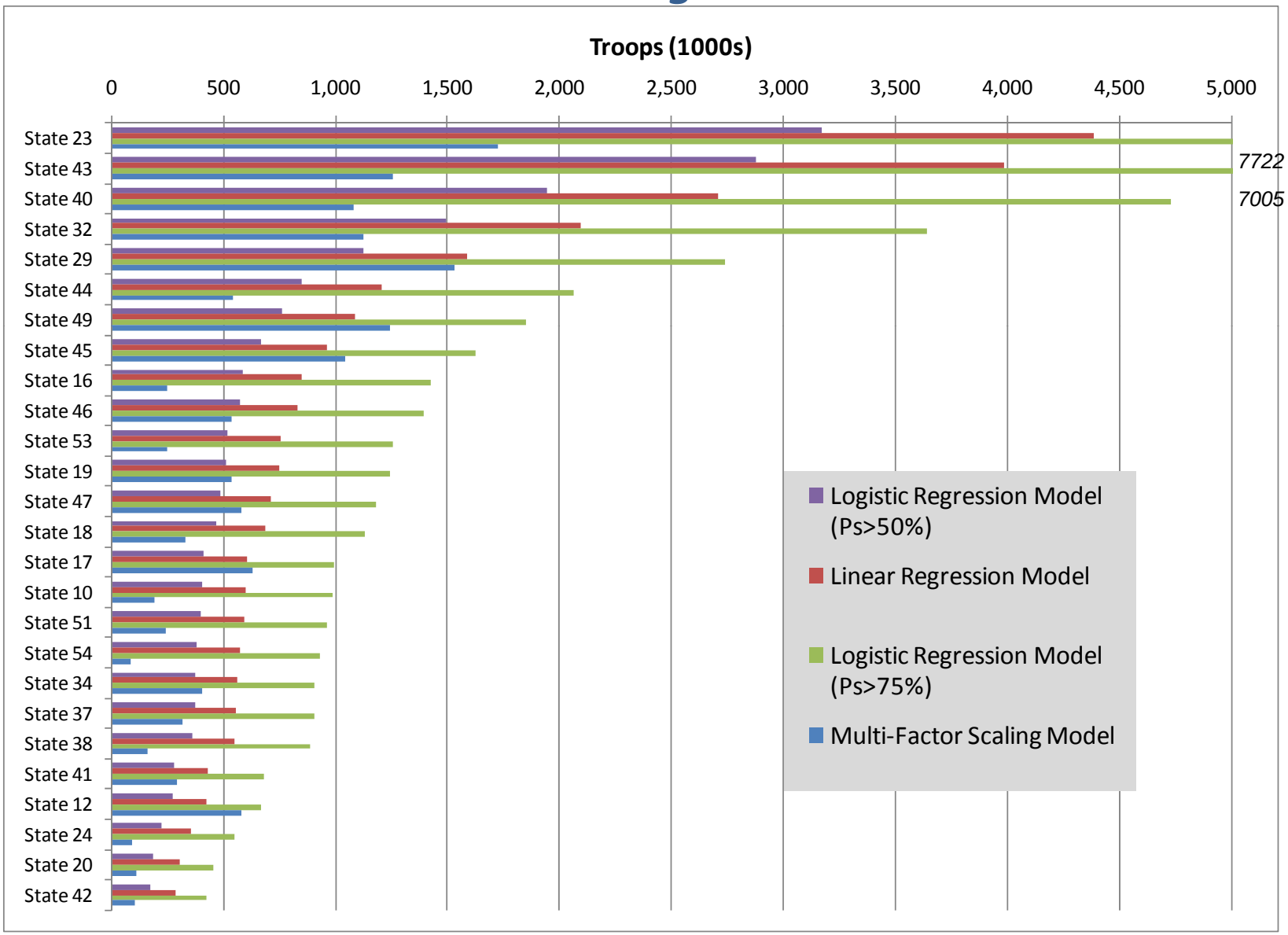
Average Troops per thousand range from less than 15 to almost 40

This chart shows the results of applying the force-sizing methodologies developed by the study to the 26 largest countries from the full set of 54 considered. The version of this chart in the Classified Appendix contains the names of the countries and is, of course, of greater interest. This unclassified version, however, illustrates the range of results obtained.

The wide range of results, depending on the methodology and assumptions used, while perhaps disconcerting to those looking for “the right answer,” is regrettably inherent in quantitative predictions of this nature. As our “baseline” estimate we have chosen the linear regression approach, since it provides a mid-range estimate with reasonably good analytical support.

These projections are for the total Blue force, which could be composed of indigenous forces and foreign or intervention forces. Furthermore, the intervention force could be composed of a mix of U.S. and allied or partner nation forces. The next step in our analysis is to estimate the breakdown of the Blue force into those components so that estimates of U.S. force requirements can be obtained.

Comparison of Force Size Projection Methodologies
Selected Larger Countries



This chart summarizes advantages and drawbacks of the three methodologies that were developed to estimate the size on the Blue force for projected future stability operations.

Comparison of Projection Methods

Methodology	*Advantages*	*Drawbacks*
Multi-Factor Scaling Model	• Includes subjective factors on country stability and cooperativeness • Does *not* depend on historical conflicts with little apparent relevance to future U.S. COIN operations	• Subjective factors lack analytic rigor • Some factors possibly redundant
Logistic Regression	Scales only on factor(s) demonstrating statistically significant relationship to troop density	• Good data not extensive • Statistical significance not robust
Linear Regressions	Sizes forces for success based only on historical "no-lose" cases	• Weak R^2 • Inclusion of constant term distorts results for smaller conflicts • Less rigorous than logistic regression

The next step, as already mentioned, is to estimate the composition of the projected Blue force in terms of indigenous, U.S. and allied/partner force components.

Projecting U.S. Force Requirements for Stability Operations

- Indigenous forces
- Allied/Partner forces

I. The Role of Indigenous Forces—Insights from Recent COIN Operations

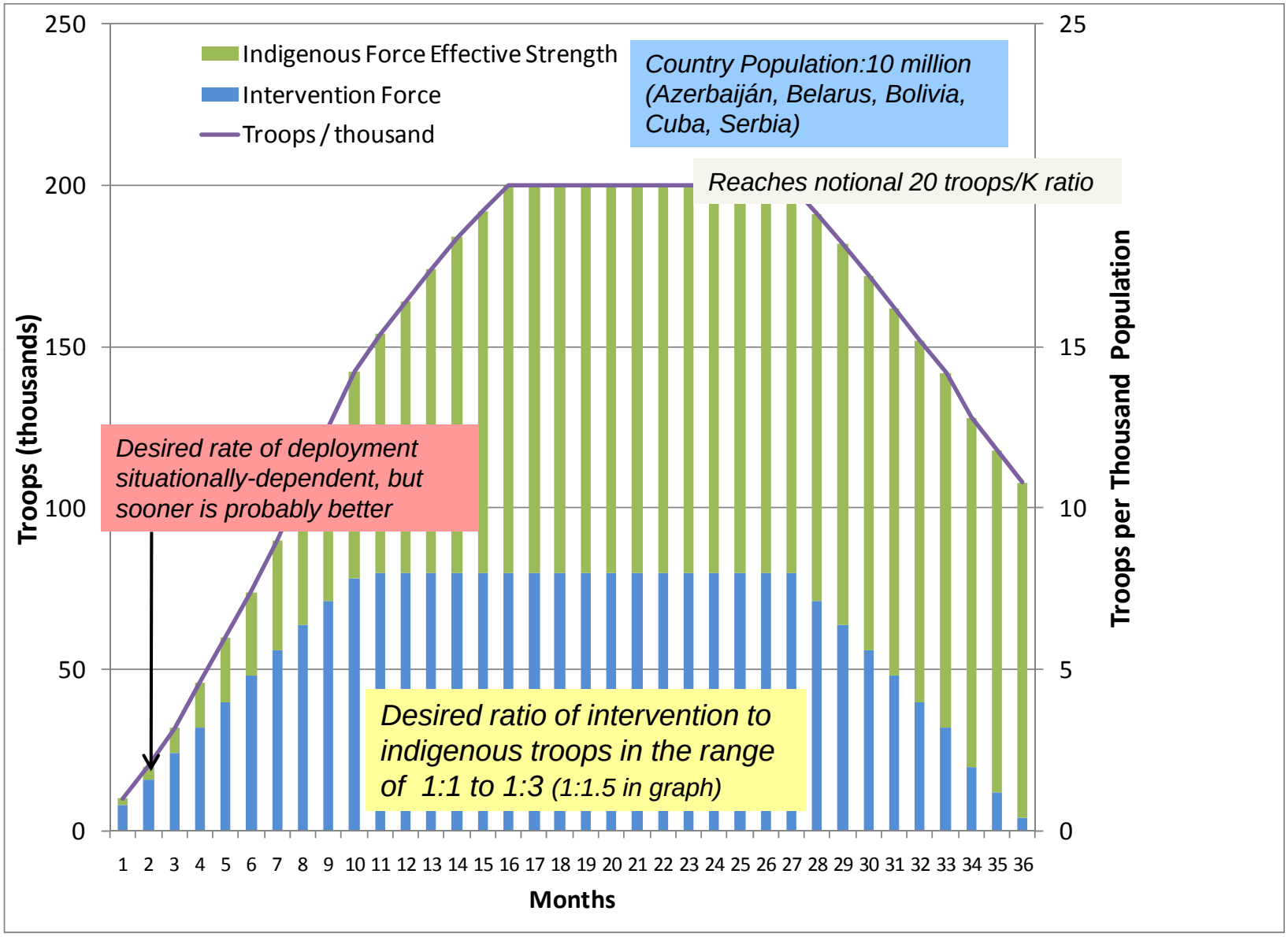
This chart, which is purely notional, illustrates the thinking that led to an approach to determine the share of a stability force that might reasonably be assumed to be satisfied by indigenous forces.

First we are only interested in situations where intervention by a foreign force is believed to be needed to achieve stability. At the beginning of the stability operation, we assume (as a limiting case for this hypothetical construct) that no indigenous forces will be capable of helping to maintain stability in the area of the stability operation. Consistent with the experience of intervention forces for a number of conflicts, the initial tasks to be performed by the intervention force are to (1) achieve a modicum of security for a base and key facilities, and (2) begin recruiting and training an indigenous force to eventually assume the role of ensuring stability. Thus this idealization shows a buildup of the intervention force as the means for entry into the host country and facilities for supporting that force are established.

Within a few months, some of the indigenous troops achieve a level of competence that enables them to perform some of the less demanding stability tasks. Within a year or so, something like a steady state is achieved in which the operations to achieve and maintain stability are shared between the intervention and indigenous forces, during which the competency of the indigenous troops continues to grow. Eventually, if the stability operation is successful, the intervention force can be withdrawn and the indigenous force scaled down to a peacetime level to maintain stability.

Under this conceptual model, then, what percentage can indigenous forces comprise of the ultimate stability force? Several charts that follow address that issue.

Notional Construct: Troop Strengths in Stability Operation



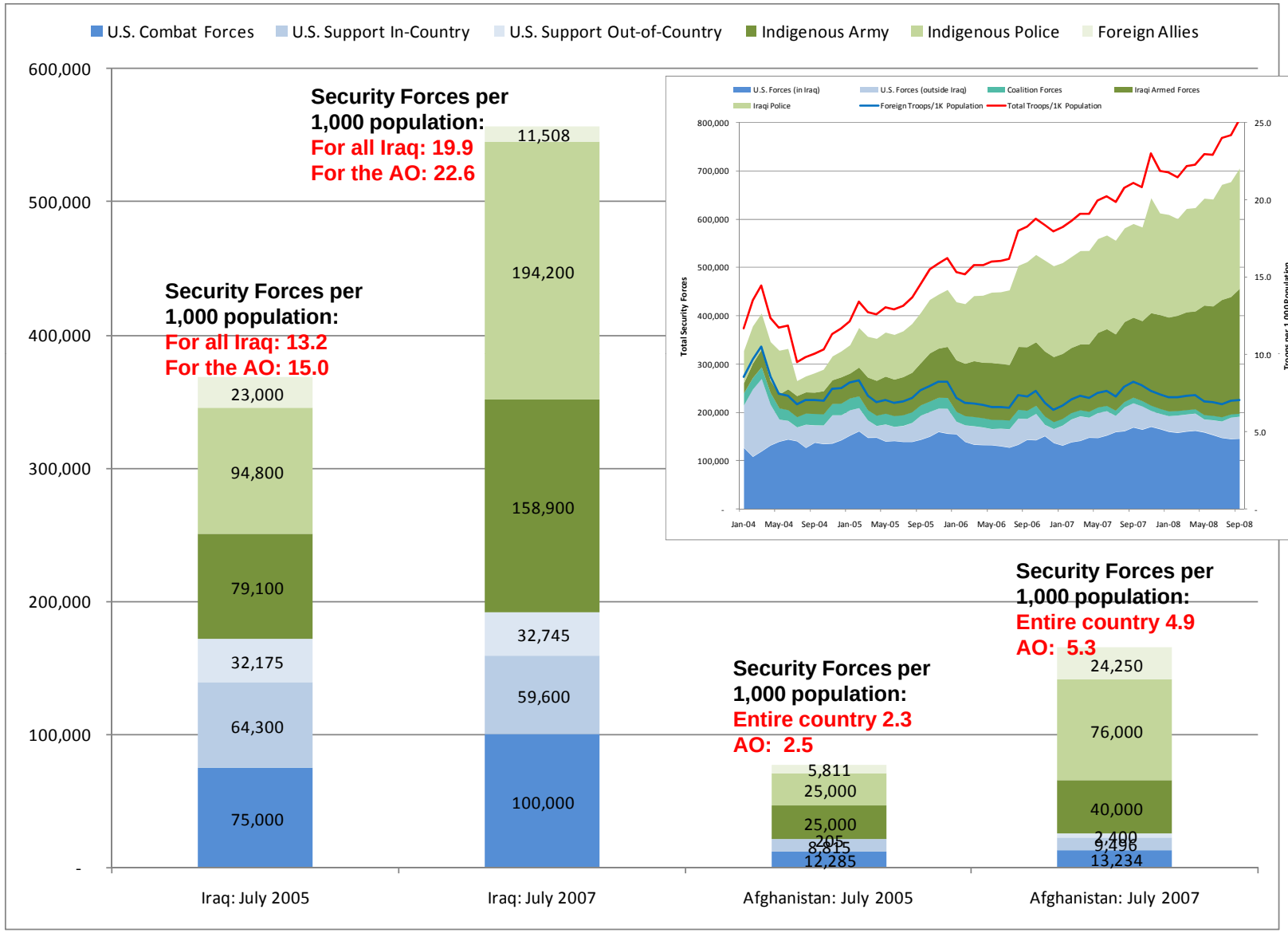
To approach the issue of the role of indigenous forces in stability operations we drew on U.S. experience in Iraq and Afghanistan. We believe these conflicts offer the best model of a COIN-focused stability operation in which U.S. forces might participate in the future. In addition, a wealth of data is available from U.S. Iraq and Afghanistan experience. Although some might charge that such an approach "re-fights the last war," the U.S. military has learned a tremendous amount as a result of its failures and then its successes in Iraq. It is now in the midst of applying that knowledge to Afghanistan. These are the wars that will shape future U.S. COIN operations.

The chart below contains summaries of the level and mix of U.S., partner/allied, and indigenous forces (both military and police) for Iraq and Afghanistan. The main chart provides figures for key points in the Iraqi and Afghan conflicts. The inset chart provides the full time series of troop strength data for Iraq.

It is interesting to note that the success of the "surge" in Iraq coincided with the achievement of a force density that exceeded the recommended minimum of 20 troops per 1,000 inhabitants in the AO. Significantly, the achievement of this density was driven by an increase in the number of indigenous forces, even though there was an increase in coalition forces as well. Clearly, improvements in strategy and tactics also played an essential role in the surge. However, we would be remiss not to observe that additional boots on the ground may have facilitated the use of those new strategies and tactics.

As a result of the additional deployments ordered in late 2009, the force density ratio in Afghanistan will begin to approach the recommended minimum. It remains to be seen what effect this will have on operations.

Total Security Forces by Type in Iraq and Afghanistan



This chart summarizes insights applicable to this study from the analysis of ongoing operations in Iraq and Afghanistan. First, it is certainly noteworthy that at the peak of the surge in Iraq the force density achieved by the combination of U.S., partners, and reasonably capable indigenous forces happened to be 20 per thousand. It is, or course, coincidental that this figure corresponds to the low end of the figure cited in FM 3-24 as the minimum threshold for success in COIN operations. However, it should be noted that the population used for the 19.9 force density figure in the chart below is the total estimated population for Iraq, not the actual "area of operations," which was considerably smaller. Based on the reduction for Iraq determined by the study, the force density would be 22.6. However, it is likely that the actual AO for the surge was smaller, meaning the effective force density was larger. (We were unable to determine the exact actual AO for the surge from the available data.)

The next insight that we drew from current operations concerns the ratio of intervention to indigenous troops in the stability force. As stated in the chart, we observed ratios in the range of 1:1 to 1:2 for the overall Iraqi stability force, corresponding in percentage terms to 50% to 67% indigenous troops.

We also obtained data on several localized operations in Iraq (to be discussed in the chart following) and held discussions with several people with direct experience in working with indigenous forces in both Iraq and Afghanistan.

One particularly well-informed person (a U.S. Marine Corps colonel) told us he believed that in order to be successful, an Afghan army unit needed to have U.S. (or equivalent) troops training, advising, and assisting in the ratio of 1:3—one advisory troop for every three Afghan troops. His judgment was also informed by experience in Iraq; in particular, he cited the Al Qaim operation reviewed in the slide following.

Analysis of Recent and Current Operations

- The combination of U.S., partner, and trained indigenous forces achieved in Iraq (in 2007) is consistent with the historically-derived rule of thumb of a minimum of 20 troops per thousand
- Ratio of intervention to indigenous forces:
  - In July 2007 in Iraq, the overall ratio of Coalition (US and partners) to indigenous troops (including police) was 1:1.7 (and continues to increase); in Afghanistan it was 1:2.3
  - A ratio between 1:1 and 1:3 was observed in several successful operations in Iraq
    - And is consistent, *on average,* with a well-informed judgment based on experience in both Afghanistan and Iraq
- Demands for combat forces and support forces have been roughly equal in terms of total personnel deployed
- Recent experience further strengthens the caveat that strategy and tactics are at least as important as force size in achieving success

This chart summarizes results from three significant operations in Iraq. These examples reinforce our findings that force densities greater than 20 in the AO should be considered necessary (but not sufficient) for success. In addition, they provide some insight into the appropriate ratio of U.S. to trained indigenous troops.

Examples of Iraq Troop-Population Ratios

A few examples of U.S. experiences in Iraq during 2005-6 illustrate the importance of force size and density in stabilizing urban areas.

	Ramadi	Tal Afar	Al Qaim
Time Frame	Winter 2005	Fall 2005	Winter-Spring 2006
City population	400,000	150,000	200,000
U.S. units	2 infantry battalions (1 USA, 1 USMC)	2 battalions plus (1 cavalry, 1 infantry, SF)	1 infantry battalion
Estimated U.S. troop count	1,800	2,200	1,200
Estimated Iraqi troop/police count	1,000	2,000	2,000
Troops / 1,000 population	7 (~65% U.S.)	29 (~50% U.S.)	16 (~40% U.S.)
Change in enemy attack levels	Up (Classified)	Down 54%	Down >70%
Participant comment	"The biggest problem was that we were so undermanned that we couldn't give the people confidence that we'd be around. As soon as you're gone, you can count on the insurgents to show up and intimidate them or punish them in retribution for their cooperation with the Coalition." – infantry company commander	"[S]aturating the area with forces is guaranteed to have a major effect. . . [The regiment] had a pretty sizable footprint in the city, and this accounts for a lot of the improvement in security and stability." – cavalry squadron commander	"At the point in time immediately following kinetic operations, a surge of forces was required, and the forces needed to be active and visible to the population." – infantry battalion commander

This chart shows our “bottom line” choice of a ratio of 40%/60% intervention to indigenous troops as a reasonably supportable planning factor for the split between intervention and indigenous troops.

A higher proportion of intervention troops is not necessarily beneficial, since indigenous troops would in general have greater cultural awareness and could help overcome communications barriers.

Ratio of Intervention to Indigenous Forces for Effective Stability Operations

- U.S. experience in Iraq and Afghanistan indicates a ratio of intervention to indigenous forces of between 1:1 and 1:3, or in terms of percentages, between 50/50% and 25/75%
- In the CAA historical sample, when the percentage of foreign troops was not either 0% or 100%, the median was 53%

Although these data and information are obviously severely limited, we believe that it is of high enough quality to provide basis for suggesting a ratio of intervention to indigenous forces for program planning purposes. The study team chose a reasonably conservative ratio of 2:3, or 40% intervention and 60% indigenous for our projections.

J. Allied/Partner Participation

Since the ultimate objective of the study is to estimate the required sizes for *U.S.* forces, the next issue that needs to be addressed is what to assume about the mix of U.S. versus allied/partner forces when an intervention force is required for a stability operation.

Of course, at the most conservative end, one can assume *no* allied or partner help. That might be the appropriate assumption in many contexts; but for programmatic force planning it may be more appropriate to make an assumption more in line with historical experience. For almost all the historical stability operations in which U.S. forces have participated, allied or partner forces have also participated.

This chart cites recent experience in Iraq and Afghanistan and one historical operation (Bosnia). Because of limited study resources we did not search for more extensive historical data on the mix. Instead our projections relied on subjective estimates of the expected allied/partner participation that were made for the 2005-6 study. Those estimates considered the factors described in the slide, and ranged from a low of 10% (allied/partner participation) to a high of 75%, with an average of 35% over the 54 countries considered.

Allied/Partner Participation

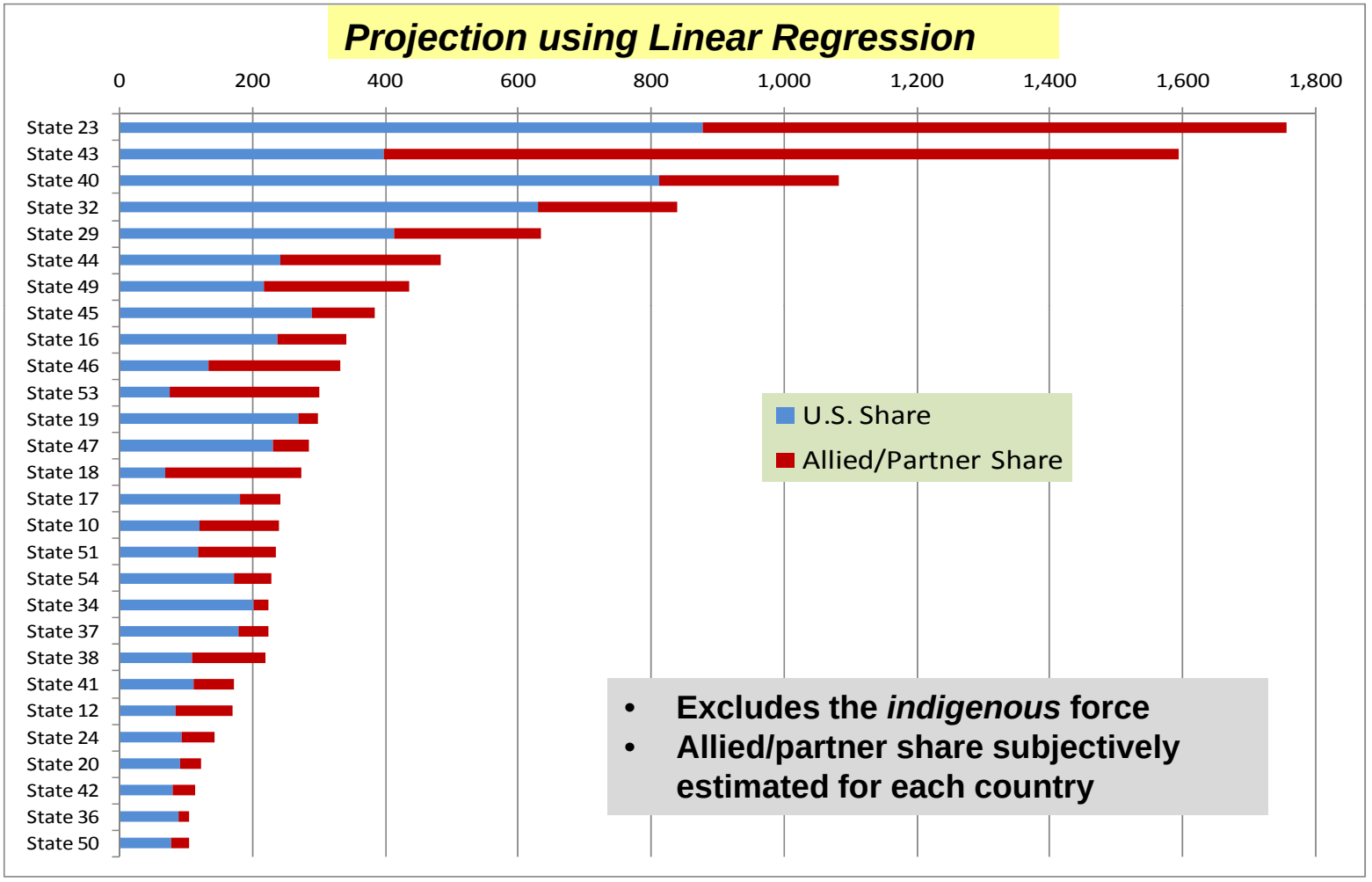
- Most actual stability operations in which U.S. forces have participated have also had significant allied or partner national troops participation
- Some salient examples of the allied/partner percentage of the intervention force
  - Iraq-- Average (Jan-04 to Sep-08): 11%; range: 5-18%
  - Bosnia--70% (at peak U.S. buildup)
  - Afghanistan-- Average: 50%; range: 42-58%
- The study used subjective factors developed in the 2006 study
  - Historical relationships
  - Relative importance to national interests
  - Geographic proximity

K. Force Size Projections

This chart displays the results of the estimates for the size of the intervention forces for a stability operation in a wide range of countries, showing the subjective breakout between U.S. and allied/partner troops. The linear regression methodology is used. The identities of the countries are given in the Classified Appendix

Obviously this approach to determining the sharing of the burden between U.S. and allied/partner forces is less than satisfying from an analytical perspective. Other approaches that have been suggested or used are to simply assume a constant share for allied/partner force participation—say 15%, or to cap their participation at some arbitrary size limit, say 100,000 troops. The best approach from an analytical perspective is to assume *no* allied/partner participation, which establishes an upper bound on U.S. required troop levels. However, use of such an upper bound for force planning would mean substantial increases in programmed force levels, which is inconsistent with resource limitations. Faced with these unpalatable alternatives, we chose to stay with the subjective factors that were used in 2005.

Force Size Projections, *Intervention* Force (US & Allies)

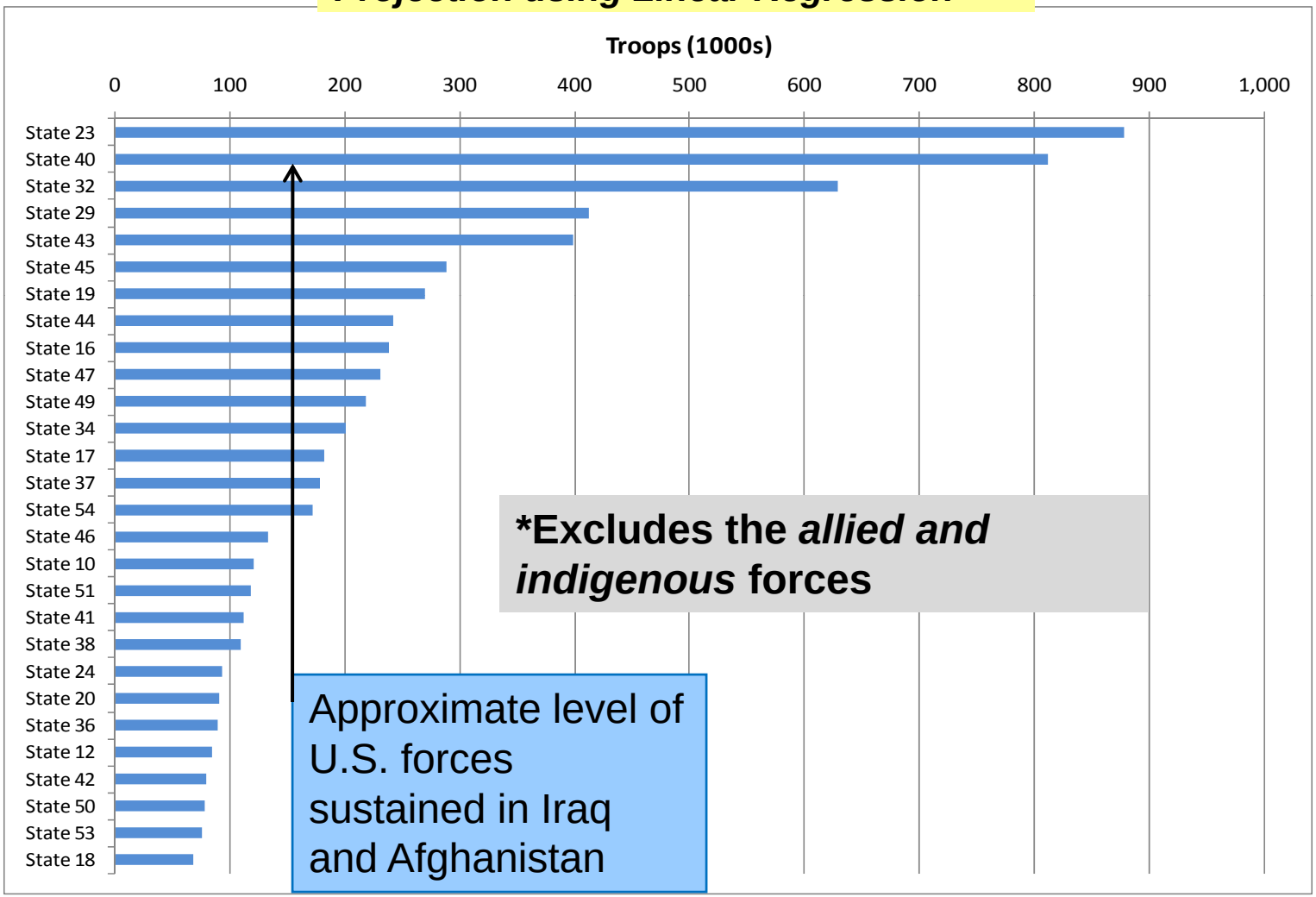


This chart is our “bottom line” estimate on U.S. forces required to perform a larger-scale stability/COIN operation in a range of countries. In the 2005-6 study a similar chart also had a overlay box showing a range of the estimated *availability* of U.S. ground forces (in terms of brigade combat teams) to sustain such an operation. As a point of reference, the U.S. has, over the past six years, sustained forces for stability operations in Iraq and Afghanistan in the 200,000 range, albeit with considerable stress on the force. Assuming that level of commitment could (and would) be sustained in a future operation, eleven of the 54 countries considered would be problematic for a U.S.-led intervention. See the Classified Appendix for the identity of the countries.

The Classified Appendix also contains projections made using our methodologies for some of the scenarios in the Defense Planning Scenario set.

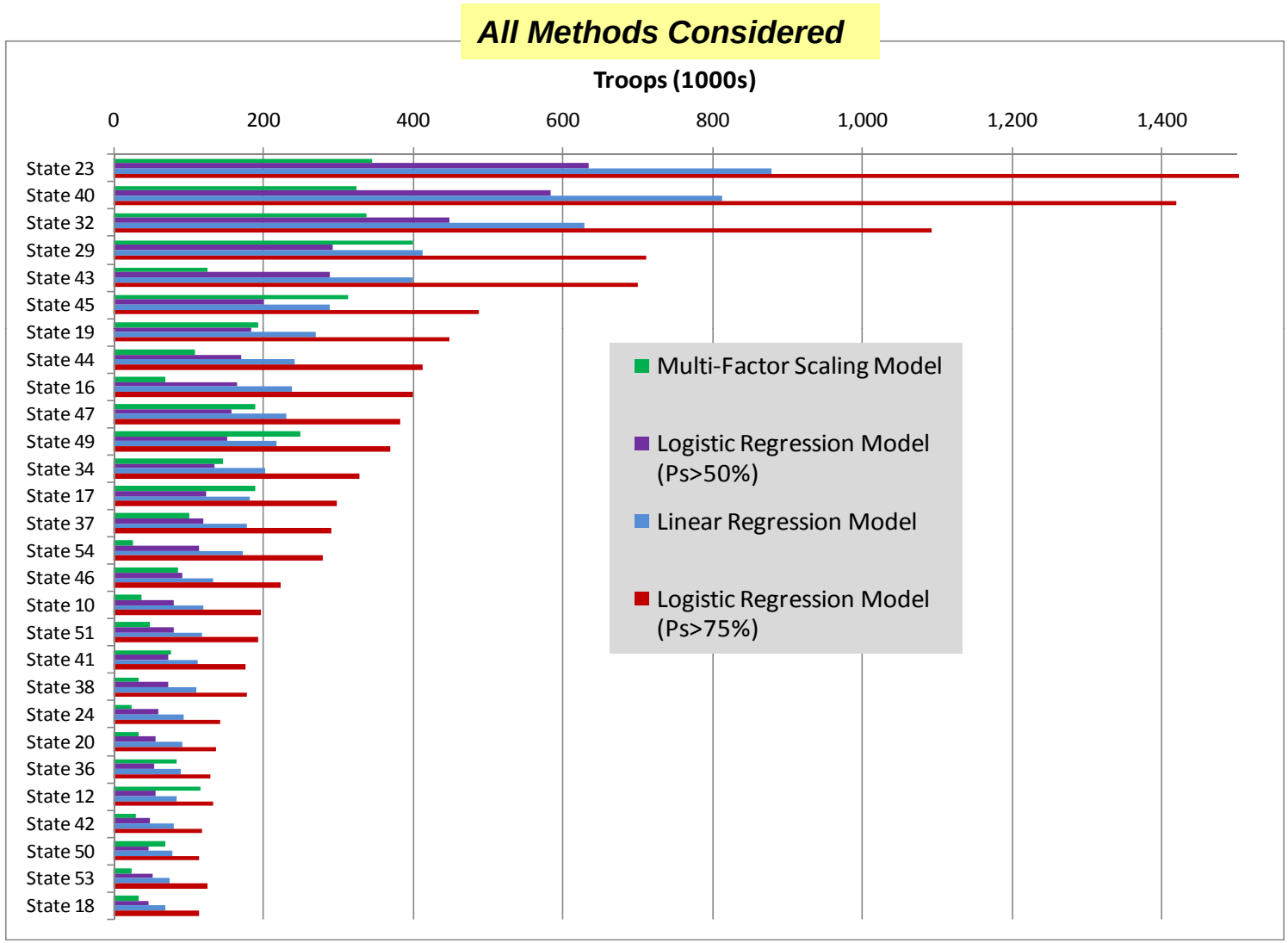
Force Size Projections, U.S. Share*

Projection using Linear Regression



This chart displays U.S. force size projections for all four of the methodologies used in the study.

Force Size Projections, U.S. Share



APPENDIX A.
HISTORICAL DATABASE AND SCORING OF OUTCOMES FOR THE LOGISTIC REGRESSION

Database for Historical Analysis

(Basic data from Center for Army Analysis Irregular Warfare Database)

Case Number	Conflict (Years)	COIN Force Peak Strength	TOTAL CI Peak Year	National Population (CI Peak Year)	Unadjusted Troops Per 1,000	Adjusted population*	Adjusted Troops Per 1,000*	IDA Scoring*
1	Aden (1963-1967)	32,500	1966	650,000	50.00	650,000	50.00	Loss
77	Afghanistan (1979-1989)	252,150	1987	15,220,000	16.57	8,899,685	28.33	Loss
4	Algeria (1954-1962)	628,238	1961	11,007,000	57.08	11,007,000	57.08	Indecisive
109	Algeria (1992-1999)	303,200	1999	30,072,000	10.08	26,469,796	11.45	Win
5	Angola (1961-1974)	67,992	1973	6,494,000	10.47	2,126,070	31.98	Indecisive
6	Angola (1975-1988)	175,450	1988	10,000,000	17.55	6,620,000	26.50	Win
7	Argentina (1969-1983)	228,500	1981	28,522,000	8.01	20,046,518	11.40	Win
10	Cambodia (1978-1989)	260,000	1987	8,738,000	29.76	1,875,977	138.59	Indecisive
11	Cameroun (1955-1960)	5,360	1958	5,191,000	1.03	2,595,500	2.07	Loss
21	Chechnya (1994-1996)	60,000	1995	1,105,500	54.27	1,105,500	54.27	Loss
55	Chechnya (1999-2004)	99,992	2000	1,105500	90.45	862,000	116.0	Win
15	Colombia (1964-2009)	306,000	2006	45,558,000	6.72	7,668,500	39.90	Win
17	Cuba (1956-1959)	30,000	1956	6,513,000	4.61	3,158,653	9.50	Loss
18	Cyprus (1955-1959)	42,278	1959	567,000	74.56	419,000	100.90	Win
19	Dhofar (1965-1976)	13,500	1975	917,000	14.72	50,000	270.00	Win
31	East Timor (1975-1999)	30,603	1999	809,000	37.83	809,000	37.83	Indecisive
20	El Salvador (1979-1992)	82,708	1988	5,032,000	16.44	2,959,522	27.95	Win
24	Greece (1946-1949)	230,612	1949	7,566,000	30.48	5,233,768	44.06	Win
25	Guatemala (1960-1996)	67,767	1992	9,745,000	6.95	3,119,000	21.73	Win
50	Guinea-Bissau (1963-1974)	32,035	1973	615,000	52.09	615,000	52.09	Loss
59	Hungary (1956)	168,500	1956	9,870,000	17.07	9,870,000	17.07	Win
29	Indochina (1946-1954)	448,241	1954	29,438,000	15.23	29,438,000	15.23	Loss
30	Indonesia (1945-1949)	180,000	1948	79,538,000	2.26	79,538,000	2.26	Loss
36	Kenya (1952-1956)	71,582	1955	6,984,000	10.25	1,500,000	47.72	Win
37	Malaya (1948-1960)	105,891	1953	5,506,000	19.23	2,750,000	38.51	Win
114	Morocco (1953-1956)	137,900	1956	10,430,800	13.22	10,430,800	13.22	Loss
38	Mozambique (1964-1974)	51,463	1973	8,755,000	5.88	1,442,639	35.67	Indecisive
39	Mozambique (1976-1992)	81,690	1990	14,200,000	5.75	6,676,929	12.23	Win
41	N. Ireland (1968-1998)	48,341	1972	1,536,500	31.46	750,000	64.45	Win
40	Namibia (1966-1989)	37,500	1985	1,131,000	33.16	657,915	57.00	Indecisive
116	Nepal (1996-2008)	162,320	2007	28,215,600	5.75	22,710,117	7.15	Loss
54	Nicaragua (1967-1979)	16,700	1979	2,451,000	6.81	1,225,265	13.63	Loss
16	Nicaragua (1981-1990)	90,300	1989	3,676,000	24.56	1,118,625	80.72	Indecisive
58	Peru (1980-1999)	550,000	1994	23,460,000	23.44	12,707,352	43.28	Win
112	Philippines (1946-1954)	56,963	1952	21,533,000	2.65	3,571,212	15.95	Win
53	Rhodesia (1972-1979)	44,790	1979	7,027,000	6.37	5,746,681	7.79	Loss
60	Sri Lanka (1983-2002)	205,300	1995	18,872,000	10.88	3,396,960	60.44	Win
119	Tunisia (1952-1956)	53,635	1954	3,794,000	14.14	3,794,000	14.14	Loss
64	Uganda (1979-1986)	41,000	1984	14,801,000	2.77	8,167,257	5.02	Loss
80	Vietnam (1965-1973)	1,637,037	1969	17,914,000	91.38	17,914,000	91.38	Indecisive
81	Yemen (1962-1970)	62,000	1965	4,000,000	15.50	4,000,000	15.50	Indecisive

* Indicates data added by IDA

SCORING OF CONFLICT OUTCOMES

Brief explanations of IDA's outcome scoring decisions for each of the 41 historical campaigns in our data set are listed in alphabetical order. The entries below also provide brief explanations of why there is disagreement among researchers on the scoring of certain campaigns' outcomes. Before turning to the specific campaigns, it is necessary to consider the definitions of success applied by different researchers.

Jason Lyall of Yale and LTC Isaiah Wilson of West Point write that "a win occurs when the insurgency is militarily defeated and its organization destroyed, or the war ends without any political concessions granted to insurgent forces." Conversely, Lyall and Wilson define a loss as "a situation in which the incumbent unilaterally conceded to all, or nearly all, insurgent demands." In between victory and defeat lie draws, which occur "when an incumbent is forced to concede to some, but not all, insurgent demands, and neither side obtains its maximal aims."[1] This standard for outcome scoring reflects simple common sense, yet may be extremely hard to apply in practice, given the unusual ways in which insurgencies end. For example, how stringent should one be in applying the standard that victory only occurs in the absence of *any* political concession? Similarly, how does one know when an incumbent has conceded to *nearly all* of the insurgents' demands? The shades of meaning in common words such as *any* and *all* are deep enough to generate persistent debate about where to draw the line between decisive outcomes and ambiguous ones.

To clarify the distinction between victory, defeat, and indecision, Andrew Hossack of the Defence Science and Technology Laboratory of the United Kingdom Ministry of Defence (DSTL) distinguishes between the military and political outcomes of an insurgency. Whereas military victory entails the possession "of an effective monopoly on any capacity to resume violence in [the] future," political success should be "judged in terms of the extent to which each player's reported initial strategic political/military objectives were met."[2] Hossack avoids the use of problematic words such as *any* and *all* to describe the threshold at which either side achieves political victory. Yet no substitute is offered. Thus, there is no way to resolve the persistent debates about whether any given

---

1 Jason Lyall and Isaiah Wilson, "Rage Against the Machines: Explaining Outcomes in Counterinsurgency Wars," *International Organization* 63, Winter 2009, 71-72.

2 Andrew Hossack and Karthik Sivansankaran, "Success Factors in CT/COIN Campaigns: Preliminary Results Arising from Current Research," Paper presented at the Cornwallis X conference, 21-24 March 200511-12.

campaign had a decisive political outcome. Although quite sensible, the distinction between military and political outcomes does not significantly reduce the subjectivity of outcome scorings, since it is hard to know at what point military operations cross the threshold from indecisive to successful. In addition, some analysts may argue that the separation of military and political objectives represents a fundamental misunderstanding of COIN, which is an inherently political form of warfare.

A different approach to the problem of outcomes is to allow subject matter experts to assess the outcome of individual conflicts, or at least to consult the publications of leading scholars. The Center for Army Analysis (CAA), IDA and the Rand Corporation (RAND) all adopted this approach in the course of their research. However, this only evades the question of what standard the scholars and experts should apply. Today, thirty-five years after the fall of Saigon, experts and non-experts alike continue to debate whether the war in Vietnam amounted to a victory, a defeat, or something in between. Although few historical debates have the political and emotional salience, for Americans, of the one about Vietnam, the outcomes of many other wars are equally ambiguous.

One might argue that the entire effort to label insurgency outcomes as a victory, defeat, or indecisive is fundamentally misguided, since these terms themselves are inherently subjective. Yet any research design that relies on ordinal logistic regression depends on the clear labeling of outcomes, since the application of logistic regression to a given data set requires the presence of a discrete outcome variable. With the exception of RAND, all of the organizations listed above rely primarily on logistic regression to conduct statistical analysis of insurgency outcomes. To circumvent the subjectivity of outcome scorings, they adjust their techniques in a sensible manner. CAA chose to exclude indecisive outcomes from its analysis. IDA and Lyall-Wilson applied logistic regression techniques twice, once counting indecisive outcomes as successful, once counting them as failures. DSTL used both military and political success as its outcome variables. As noted in the main body of this briefing, none of these approaches generated results that were sufficiently robust to overcome other researchers' objections to the techniques on which they were based. Given the relatively small size of the data sets involved, disagreements about the scoring of several key outcomes resulted in completely divergent findings about the significance of leading independent variables.

As a result, participants in the December 14 seminar on force sizing recommended that future studies focus on sub-national data from individual conflicts,

rather than conducting further cross-national studies that seek to identify single variables that influence insurgency outcomes across the globe.

DISCUSSION OF OUTCOME SCORING BY CONFLICT

Aden (1963-1967) – Failure. There is full agreement that the British failed to suppress the insurgency in Aden. Although the British had considerable manpower, the political situation was highly unfavorable. Nationalist insurgents took power shortly after the British departure.

Afghanistan (1979-1989) – Failure. The Soviet Union persistently failed to control rural areas. It ultimately withdrew from Afghanistan in 1989 amidst mounting casualties. The Soviet-installed communist government fell to an insurgent offensive three years after the Soviet withdrawal. One organization coded the outcome of the Soviet campaign as indecisive because the insurgents failed to best Soviet forces on the battlefield and the Soviet-installed government remained in place far longer than contemporary observers expected.

Algeria I (1954-1962) – Indecisive. In the later years of the war, French forces significantly curtailed the insurgents' ability to conduct operations and threaten the population's security. However, as a result of intense controversy both at home and abroad, the French abandoned their initial objective of maintaining Algeria as an integral part of France. For this reason, the majority of organizations code the outcome of this campaign as a failure.

Algeria II (1992-1999) – Success. After a long and brutal conflict, Algeria's secular authoritarian regime crushed the Islamist opposition. One organization coded the outcome of this campaign as indecisive because some terrorist activity persists in Algeria.

Angola I (1961-1974) – Indecisive. This case is similar to **Mozambique I**. Portuguese forces reversed the insurgents' momentum, sowed conflict among insurgent movements, and secured all but the most remote areas of their colony in Angola. However, after the fall of Portugal's right-wing authoritarian regime in 1974, the new left-of-center government in Lisbon chose to grant independence to all of its colonies in Africa. This outcome illustrates how political events tangential to the war can reverse its outcome. Because of the influence of exogenous events, IDA coded this outcome as indecisive, although it is the only organization that did so.

Angola II (1975-2002) – Success. Negotiations led to a pause in the Angolan civil war after a decade and a half of fighting. However, in spite of the presence of a U.N.

peacekeeping force and relatively fair national elections, the conflict resumed and the government prevailed. One organization codes this outcome as indecisive.

Argentina (1969-1983) – Success. After a brutal campaign that was the first to be known as a "dirty war" (*guerra sucia*), the military regime crushed the insurgents. The military regime allowed national elections and stepped down three years after the end of the insurgency, for reasons that most analysts consider unrelated. However, one organization codes this outcome as indecisive.

Cambodia (1978-1989) – Indecisive. The end of the Cold War facilitated a negotiated outcome to the Cambodian insurgency. Free elections led to the inauguration of a coalition government led by non-communist forces, although the communists gradually hollowed out the electoral system and re-established *de facto* rule. One organization codes this outcome as a failure.

Cameroon (1955-1960) – Failure. There is full agreement that the French failed to suppress the insurgency in Cameroon, resulting in independence for the colony.

Chechnya [Russia] (1994-1996) – Failure. Russian forces withdrew from Chechnya in 1996, although Chechnya remained technically a part of Russia. Two organizations code this outcome as indecisive.

Chechnya [Russia] (1999-2004) – Win. In 1999 Russian forces again fought the Chechen insurgents, employing a larger force and a better strategy based on recruiting greater numbers of Chechen troops for the COIN force. This time success was achieved in defeating the insurgency militarily.

Colombia (1964-2009) – Success. In the early 1990s, Colombia approached the brink of becoming a failed state. However, a series of military and political reforms by a new president led to a dramatic reversal in the campaign against narco-insurgent forces. The guerrillas currently maintain a foothold in remote border areas adjacent to Venezuela and Ecuador. Opinion is divided as to whether the current state of affairs is decisive.

Cuba (1956-1959) – Failure. There is full agreement that the insurgent force led by Fidel Castro defeated the government.

Cyprus (1955-1959) – Success. The insurgents ended their struggle for unification with Greece, accepting instead the outcome of negotiations between the British colonial authority and the civilian opposition. The British then granted independence to Cyprus, while retaining rights to a major military installation. Opinion is

divided as to whether Cypriot independence mitigated the British success, or whether decolonization was inevitable and the negotiated outcome amounted to success.

El Salvador (1979-1992) – Success. Initially determined to overthrow the government, the insurgents laid down their arms in exchange for the right to contest elections as a civilian party and to employ thousands of insurgents in the reformed post-war military and police organizations. A minority considers these concessions to be a mitigation of the government's success, given that they reflected the military's inability to achieve a decisive outcome on the battlefield.

Greece (1946-1949) – Success. There is full agreement that the government defeated communist insurgent forces in the years following World War II.

Guatemala (1960-1996) – Success. The military regime prevailed in Guatemala, in spite of international condemnation of its horrific human rights violations. In the closing decade of the war, the military initiated a process of gradual democratization. One organization codes this outcome as indecisive.

Guinea-Bissau (1963-1974) – Failure. There is full agreement that Portugal failed to suppress the insurgency in their colony of Guiné, known after independence as Guinea-Bissau. In contrast to **Angola I** and **Mozambique I**, Portuguese forces gradually lost control of the territory they were defending. The most important causes of the insurgents' success were the unity and skill of their leaders as well as the dense, swamp-like terrain. Although the war officially ended with the fall of the right-wing government in Lisbon in 1974, its defeat had already seemed inevitable.

Hungary (1956) – Success. There is full agreement that Soviet forces crushed the Hungarian uprising in 1956. However, some consider this conflict to be an insurrection, not an insurgency.

Indochina (1946-1954) – Failure. There is full agreement that Vietnamese forces defeated the French in Indochina, although a *de facto* partition preserved pro-Western forces' control of southern Vietnam.

Indonesia (1945-1949) – Failure. There is full agreement that insurgents defeated Dutch forces in what would become the independent state of Indonesia.

Indonesia [East Timor] (1975-1999) – Indecisive. Indonesian forces retained control of East Timor for a quarter of a century, although they were unable to eradicate insurgent forces. Indonesia granted independence to East Timor after the fall of the Suharto dictatorship, which initially occupied the island in 1975. The Suharto regime fell

for reasons unrelated to the conflict in East Timor, although the majority of organizations still code the outcome as a defeat for Indonesia.

Kenya (1952-1956) – Success. There is full agreement that the British defeated the Mau-Mau insurgency in the 1950s, although historians have been increasingly critical of British brutality.

Malaya (1948-1960) – Success. There is full agreement that the British defeated the Chinese communist insurgents in Malaya.

Morocco (1953-1956) – Failure. This case is similar to **Tunisia**. The French granted independence to Morocco. The insurgents did not take power. The conflict between the insurgents and the colonial forces was limited. A more stringent definition of insurgency and counterinsurgency might exclude this case.

Mozambique I (1964-1974) – Indecisive. This case is similar to **Angola I**. Portuguese forces secured most of their colony's territory and population, although the insurgents maintained a presence in remote areas inhabited by ethnic minorities hostile to the government. After the fall of Portugal's right-wing authoritarian regime in 1974, the new left-of-center government in Lisbon chose to grant independence to all of its colonies in Africa. This outcome illustrates how political events tangential to the war can reverse its outcome. Because of the influence of exogenous events, IDA coded this outcome as indecisive, although it is the only organization that did so.

Mozambique II (1976-1992) – Success. The negotiations that ended the Mozambican civil war in the mid-1990s allowed the insurgents to organize a political party and compete in post-war elections. There was no decisive outcome on the battlefield. Opinion is divided as to whether this outcome was decisive.

Namibia (1966-1989) – Indecisive. South African forces were able to prevent a military win by the insurgents; however, South Africa chose to withdraw under strong international pressure, concluding that the costs of the long-running conflict outweighed the benefits of maintaining the status quo. The historical literature on this conflict is sparse.

Nepal (1996-2008) – Failure. In 2005, the king of Nepal suspended parliament and took power for himself in order to confront the Maoist insurgency. The king's decision led to massive protests supported by both the Maoists and the parliamentary opposition. Ultimately, the monarchy was abolished and elections held under a new constitution, with the Maoists winning the first elections. However, the Maoists lost

control of the government shortly thereafter. Opinion is divided as to whether this outcome is indecisive, or represents a failure for the government.

Nicaragua I (1967-1979) – Failure. In 1979, Sandinista insurgents consolidated their control of the country after the Somoza dictatorships and its armed forces disintegrated. One organization codes this outcome as indecisive, possibly because the Sandinistas promised to hold free elections and appointed two moderates to the interim junta. (Free elections were never held and the moderates were soon marginalized.)

Nicaragua II (1981-1990) – Indecisive. After supporting the Contra insurgents for several years, the U.S. government lent its support to negotiations, which resulted in an agreement to hold free elections. Initially opposed to free elections as a matter of principle, the Sandinistas calculated that it could easily defeat their civilian opposition. In a stunning upset, the opposition prevailed, putting an end to the regime. A minority consider this outcome to be a success for the Sandinista government.

Northern Ireland [UK] (1968-1998) – Success. After three decades of conflict, a peace accord was negotiated. A minority of organizations consider the length of the conflict and the need to negotiate with the insurgents as indications that the outcome was not decisive.

Oman [Dhofar] (1965-1976) – Success. There is full agreement that Omani forces defeated the insurgents in Dhofar after British advisers formulated a new strategy that focused on winning the support of the population.

Peru (1980-1999) – Success. The Peruvian military recovered from the brink of defeat in the late 1980s. It then implemented a new strategy and decapitated the leadership of the Shining Path. One organization codes this outcome as indecisive, because low-level guerrilla violence began to recur several years after the military's initial success and continues to the present day.

Philippines (1946-1954) – Success. There is full agreement that the Filipino government prevailed over the Huk insurgents in the mid-1950s, after Secretary of Defense (later President) Ramon Magsaysay implemented wide-ranging reforms designed to win popular support.

Rhodesia (1972-1979) – Failure. The white minority government negotiated an agreement that paved the way for majority (black) rule. One organization codes this outcome as a success for the government, although this may be the result of a typographical error.

South Vietnam (1965-1973) – Indecisive. Historians continue to debate whether the United States failed in Vietnam, or whether its revised counterinsurgency strategy effectively shut down the insurgency, forcing North Vietnam to launch a conventional invasion. The majority of organizations code the outcome of this campaign as a failure.

Sri Lanka (1983-2009) – Success. The Sri Lankan government decisively defeated the Tamil insurgency in 2009. There is considerable disagreement, however, about when the conflict began and whether it should be thought of as multiple conflicts. Thus, two out of four organizations code the outcome as indecisive.

Tunisia (1952-1956) – Failure. This case is similar to **Morocco**. The French granted independence to Tunisia. The insurgents did not take power. The conflict between the insurgents and the colonial forces was limited. A more stringent definition of insurgency and counterinsurgency might exclude this case.

Uganda (1979-1986) – Failure. There is full agreement that the insurgents overthrew the government of Uganda.

Yemen (1962-1970) – Indecisive. Although nationalist forces prevailed, royalist insurgents remained influential in government after the end of hostilities. Two organizations code this outcome as asuccess for the nationalist government. One organization codes the outcome as a failure for the government, although this may be a typographical error.

APPENDIX B.
ABBREVIATIONS

AO	Area of Operation
CAA	Center for Army Analysis
CAPE	Cost Assessment and Program Evaluation
COIN	Counterinsurgency
DoD	Department of Defense
DoDI	Department of Defense Instruction
DSTL	Defence Science and Technology Laboratory (of the United Kingdom Ministry of Defence)
DPS	Defense Planning Scenario
FD	Force Density
FM	Field Manual
FSI	Failed State Index
IDA	Institute for Defense Analyses
MFSM	Multi-Factor Scaling Methodology
OA	Operations Availability
OIF	Operation Iraqi Freedom
PA&E	Program Analysis and Evaluation
QDR	Quadrennial Defense Review
WMD	Weapons of Mass Destruction

APPENDIX C.
REFERENCES

Army Field Manual 3-24/Marine Corps Warfighting Publication 3-33.5, *Counterinsurgency*, Headquarters, Department of the Army, December 2006.

Belasco, Amy. "Troop Levels in the Afghan and Iraq Wars, FY2001-FY2012: Cost and Other Potential Issues." Congressional Research Service, R40682, 2 July 2009.

Blaho, Justine and Lisa Kaiser. "A Predictive Model for Irregular Wars Based on a Structured Historical Database." Presentation at the 77th MORS Symposium, Working Group 16 – Special Operations & Irregular Warfare, Fort Leavenworth, KS, 16-18 June 2009.

Dobbins, James, Keith Crane, and Seth G. Jones. *The Beginner's Guide to Nation-Building*. Santa Monica, CA: The RAND Corporation, 2007.

Galula, David. *Counterinsurgency Warfare: Theory and Practice*. New York: Frederick A. Praeger, 1964 [2005].

Gompert, David and John Gordon IV. *War By Other Means: Building Complete and Balanced Capabilities for Counterinsurgency*. Santa Monica, CA: The RAND Corporation, 2008.

Hosmer, David W. and Stanley Lemeshow. *Applied Logistic Regression*. 2nd ed. New Jersey: John Wiley & Sons, 2000.

Hossack, Andrew and Karthik Sivasankaran. "Success Factors in CT/COIN Campaigns: Preliminary results arising from current research." Presentation at the Cornwallis Group X: Analysis for New and Emerging Societal Conflicts. The Canadian Royal Military College, Kingston, Ontario, Canada, 21-24 March 2005. Accessed 22 March 2010 at http://www.thecornwallisgroup.org/workshop_2005.php

Joint Publication 1-02, Department of Defense Dictionary of Military and Associated Terms, 12 April 2001 (as amended through 19 August 2009).

Kneece, R. Royce, Jr., Raphael Bonoan, and James S. Thomason. "Stability Operations: A Range of Possibilities (U)." IDA Document D-3231. Alexandria, VA: Institute for Defense Analyses, January 2006.

Lyall, Jason and Isaiah Wilson. "Rage Against the Machines: Explaining Outcomes in Counterinsurgency Wars." *International Organization* 63 (January2009): 61-106.

McGrath, John J. "Boots on the Ground: Force density in Contingency Operations." *Global War on Terrorism Occasional Paper* 16, Fort Leavenworth, KS: Combat Studies Institute Press, 2006.

O'Hanlon, Michael, Ian S. Livingston, and Heather L. Messera. Afghanistan Index: Tracking Progress and Security in Post 9/11 Afghanistan, 19 February 2010. Accessed 22 March 2010 at www.brookings.edu/afghanistanindex.

O'Hanlon, Michael and Ian Livingston, Iraq Index: Tracking Variables of Reconstruction & Security in Post-Saddam Iraq, 11 March 2010. Accessed 22 March 2010 at www.brookings.edu/iraqindex.

Quinlivan, James T. "Force Requirements in Stability Operations." *Parameters* 24 no. 4: (Winter 1995-1996): 59-69.

Quinlivan, James T. "Burden of Victory: The Painful Arithmetic of Stability Operations." RAND *Review* 27, no. 2 (Summer 2003): 28-29.

U.S. Department of Defense. *Military Support for Stability, Security, Transition, and Reconstruction*. Department of Defense Directive (DoDD) 3000.05. November 28, 2005.

U.S. Department of Defense. *Irregular Warfare (IW)*. Department of Defense Directive (DoDD) 3000.07. December 1, 2008.

U.S. Department of Defense. *Stability Operations*. Department of Defense Instruction (DoDI) 3000.05. September 16, 2009.

REPORT DOCUMENTATION PAGE		Form Approved OMB No. 0704-0188

1. REPORT DATE (DD-MM-YY)	2. REPORT TYPE	3. DATES COVERED (From – To)
March 2010	Final	

4. TITLE AND SUBTITLE	5a. CONTRACT NO.
Force Sizing for Stability Operations Main Report	DASW01-04-C-0003
	5b. GRANT NO.
	5c. PROGRAM ELEMENT NO(S).

6. AUTHOR(S)	5d. PROJECT NO.
R. Royce Kneece, Jr., David A. Adesnik, Jason A. Dechant, Michael F. Fitzsimmons, Arthur Fries, Mark E. Tillman	
	5e. TASK NO. BA-6-3005
	5f. WORK UNIT NO.

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)	8. PERFORMING ORGANIZATION REPORT NO.
Institute for Defense Analyses 4850 Mark Center Drive Alexandria, VA 22311-1882	IDA Paper P-4556, Main Body

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)	10. SPONSOR'S / MONITOR'S ACRONYM(S)
Director, Cost Assessment and Program Evaluation, OD (CAPE) Office of Secretary of Defense Pentagon Washington, DC 20301	CAPE
	11. SPONSOR'S / MONITOR'S REPORT NO(S).

12. DISTRIBUTION / AVAILABILITY STATEMENT

Approved for public release; distribution unlimited.

13. SUPPLEMENTARY NOTES

14. The objective of this task was to update and extend previous IDA work in developing force sizing parameters for stability operations for use in program planning. The study investigated previous research and data sets related to force sizing for stability operations, and adopted for further analysis a historical data set constructed by the Center for Army Analysis covering over post-WWII 100 contingencies involving counterinsurgency and peacekeeping operations. Using a selected subset of those data, the study conducted statistical analyses to determine statistically significant factors leading to success. This analysis allowed IDA to establish relationships between "force density," defined as troops per 1000 of population, and success in counterinsurgency and similar operations. These relationships, along with data from more current operations in Iraq and Afghanistan, were used to develop several methodologies that provide a range of estimates for force sizes required for a stability operation in a wide range of countries. The estimating techniques were also applied to stability operations within the Defense Planning Scenarios.

15. SUBJECT TERMS

Stability operations, counterinsurgency operations, force sizing, irregular warfare, historical counterinsurgency conflicts

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NO. OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			19b. TELEPHONE NUMBER (Include Area Code)
U	U	U	UU	136	