Spothlight on data quality: mixing methods to ensure that data reflect reality

Center for Advanced Study in the Social Sciences

WORKING PAPERS

SPOTLIGHT ON DATA QUALITY:
MIXING METHODS TO ENSURE THAT DATA
REFLECT REALITY

Leonid Peisakhin and Paul Pinto

Instituto Juan March de Estudios e Investigaciones

SPOTLIGHT ON DATA QUALITY:
MIXING METHODS TO ENSURE THAT DATA REFLECT REALITY

Leonid Peisakhin and Paul Pinto

Leonid Peisakhin is Junior Research Fellow at the Center for Advanced Study in the Social Sciences, Juan March Institute, Madrid. Paul Pinto is a PhD candidate at the Department of Political Science, Yale University.

Abstract

Many scholars believe that multimethod research is deeply problematic because of philosophical differences underlying the qualitative, formal modeling, and quantitative approaches to research in the social sciences. In this article we advocate a non-controversial use of multiple methods by showcasing how multimethod research can be used to diagnose and remedy data problems. Issues of data quality exist prior to any philosophical disagreement about modes of analysis. We suggest that social scientists are not sufficiently focused on ensuring that the data they use are an adequate reflection of the phenomena that are being studied, and that there is insufficient appreciation for the fact that every data generation strategy comes with its own set of limitations and potential biases. We argue that using multiple methods in conjunction is an effective means of resolving these biases and illustrate our methodological prescriptions with reference to our own project on the causes of political loyalties among Chinese POWs in the Korean War.[*]

Keywords: multimethod research, data quality, political psychology, political loyalty China, prisoner of war

INTRODUCTION

A half-century long debate concerning the permissibility, practicality, and usefulness of multimethod research has been ongoing in the social sciences with scant hope of reconciliation in the near future. The fundamental philosophical disagreement is centered on differences in world-view: whereas statistically-oriented researchers are prone to see the world as a stochastic place with a lot of "noise", scholars favoring formal or qualitative methods tend to take a rather more deterministic view of social reality. Consider the role of outliers: for statisticians, data points at the tails of the distribution may be explained away as a reflection of error and noise; by contrast, outliers are either not acceptable to formal theorists within the scope of their model or are a source of additional variables for qualitative researchers. The traditional view in qualitative research contends that a causal process which is well established in a single case study or in a small number of cases is likely generalizable. Quantitatively-minded scholars dismiss such findings as invalid because of a small number of observations and ordinarily a large number of independent variables. This clash of alternative world views played itself out in political science in indirect exchanges between Sartori (1970), who claimed that large-n analysis is an exercise in "conceptual stretching", and Lijphart (1971), who argued that no reliable conclusions can be drawn from small-n analyses. Elsewhere in the social sciences the debate has been even more acrimonious. For instance, in education studies and psychology the clash between quantitative and qualitative scholars, termed "paradigm wars" (Gage 1989), left festering wounds. In sociology, these differences resulted in a spat between survey researchers and field workers (Becker and Geer 1957; Trow 1957).

Despite the persistence of philosophical disagreements and skepticism about the practical viability of single-author multimethod work because of prohibitive training costs, multimethod studies are becoming increasingly fashionable across all subfields in political science. Scholars increasingly feel that "combining methods provides opportunities for the development and testing of theories that no single method can match" (Bennett and Braumoeller 2005: 1), and there is growing appreciation for the fact that case-studies, statistics, and formal methods can be used iteratively to generate and test theories (cf. Lieberman 2005; Symposium 2007: 22-24). A marriage of methodological convenience is taking place across the social sciences: Tashakkori and Teddlie (2002) are advocating it in psychology and education studies; Brewer and Hunter (2005) are promoting the multimethod agenda in sociology; and most post-KKV (1994) methodological literature in political science has advocated various combinations of methods (Brady and Collier 2004; George and Bennett 2004; Gerring 2007). Yet we cannot say that multimethod research has become uncontroversial and mainstream. There is still a strong, and perhaps even dominant, sentiment that "the basic question of whether and exactly how qualitative and quantitative approaches can be combined still needs to be resolved. This challenge involves not only practical problems but also philosophical ones" (Mahoney 2010: 141).

In this paper we showcase an uncontroversial and underdiscussed application of multimethod research, one that bypasses the philosophical debate about different ways to sift through social reality. The starting point for our argument is an observation that in recent decades political scientists have become more concerned with methodological proficiency than with the quality of their data inputs. We have no bone to pick with methodological proficiency – a good scholar must have a sound command of her tools. Yet if underlying data quality is as important as methodological expertise, it follows that significant disciplinary returns may be reaped by investing similar rigor in building solid data foundations as is currently lavished on advanced identification techniques. This is where multimethod work comes in. We argue that multiple methods can be leveraged effectively to ensure that our data inputs are in fact fairly representative of the phenomenon under study. In making this argument we posit that issues of data

quality exist prior to debates about differences in methodological worldview. Simply put, a good scholar must ensure that his data are a decent reflection of social reality before entering into debates about the best way to draw inferences from these data. Notably, the term 'method' is commonly used to refer to two different processes: data collection (as in 'she uses the survey method') and data analysis (as in 'he is an expert in the econometric method'). In most studies, the intended mode of data analysis predetermines how the data are collected. Put simply, modes of analysis often constrain the researcher's repertoire of data gathering techniques and can in this way introduce unexpected biases into his results. In this article we use the term 'method' to refer to data collection. Also, we readily admit that what constitutes a 'decent reflection' of social reality is subject to dispute. But, while it may not be possible to ultimately know whether our data are a true reflection of the complex social phenomena that we study, we can certainly kick the tires and search for obvious flaws, such as incompleteness or willful bias of a subset of sources that we use.

We are, of course, not the first to suggest that multimethod research is an effective strategy for improving data quality. Campbell and Fiske, working in the field of psychiatry, made this point in 1959 when they noted that the same underlying psychiatric condition can look very differently depending on the data-gathering technique used to describe, code, or quantify the manifestations of that condition. Realizing that data-generation techniques sometimes introduce biases into our data, Campbell and Fiske argued that multiple methods of data gathering can be used fruitfully to ensure that the data we collect are a realistic reflection of the phenomenon under study. Campbell and Fiske's work inspired an insightful literature that further developed the rationale for and implications of data triangulation and the mixing of methods and remains highly influential in psychology, education and public health. Unfortunately, this literature has received relatively little exposure in the other social sciences and particularly in political science. This article is an attempt to shine the spotlight once again on the usefulness of the multimethod paradigm for establishing data quality.

The methodological point that we make here is directly derived from a set of experiences that we had while working on a project on the causes of political loyalties among Chinese prisoners-of-war (POWs) who found themselves in US custody during the Korean War (1950-53). In this paper we outline the evolution of this research project with a view to making our methodological argument by way of a specific example. One fascinating aspect of the POW experience in the Korean War was that prisoners on both sides were given an opportunity not to return to their homeland. While only 21 American soldiers (of 3,616 captured by the North Koreans and the Chinese)[1] decided not to return to the US, two-thirds of the 21,629 Chinese POWs held by the US and its allies opted to defect from Communist China in favor of the Republic of Taiwan, which was controlled by Chinese Nationalists under Chiang Kai-shek.[2]

Several data-gathering strategies were available to us as we set about exploring the reasons behind the Chinese POWs' repatriation decisions. We could turn to US archival sources and detailed military interrogation reports of the POWs or go directly to the source and interview the surviving POWs in China and Taiwan. For reasons that will become clear later we eventually ended up pursuing both of these research strategies. Towards the end of this fascinating, albeit lengthy, data-gathering experience it became obvious that our initial conclusions based on regression analyses of data contained in US interrogation reports (and which we almost published) were deeply flawed. Evidence from qualitative interviews made it clear that the reality which we were trying to

---

[1] The precise number of U.S. soldiers captured by the Chinese and North Korean forces is still contested. Here we use the official statistic provided by the Chinese delegate at the armistice negotiations.

[2] United Nations Command Military Armistice Commission (UNCMAC) report. Records of the US Army, Pacific. RG 550. US National Archives.

capture was considerably more complex than we had assumed after reading through the US archival sources. Data triangulation between these two quite different approaches allowed us make substantial progress towards completing the dataset and balancing the political biases on either side, laying an appropriate foundation on which to draw inferences about repatriation. The story that we tell here might look like a unique case – and the principle of voluntary repatriation has indeed been only rarely applied in the history of human conflict – yet the problem of data quality is relevant to every project in the social sciences to a varying degree. This problem is particularly pronounced in datasets where the inputs originate with an institution or set of individuals with a particularistic interest or political bias, where the data are drawn from what remains (or is allowed to remain) in the archives, or where the data are standardized and reported in a uniform fashion across nations with wildly varying social conditions and political histories. Therefore, we address this article to both quantitatively and qualitatively-minded social scientists in the hope that it will go some way towards raising general awareness about the importance of data quality.

The article proceeds as follows: First, we provide a brief introduction to our substantive project on the determinants of political loyalty among the Chinese POWs. Next, we walk the reader through the evolution of this project, reporting our empirical findings and the way we revised them as we expanded our universe of data gathering techniques and modes of data analysis. We then restate in general terms the methodological point at the core of this paper. In the conclusion we call for greater transparency on data sources and data gathering techniques.

THE HISTORICAL CONTEXT BEHIND VOLUNTARY REPATRIATION

Before we plunge into a narrative that describes the evolution of our research project it seems necessary to describe the underlying historical context. The empirical situation that gave rise to our project – the decision by warring powers to allow POWs in the Korean War to refuse repatriation – almost did not come to pass. The United States, motivated by humanitarian concern for the welfare of soldiers from authoritarian nations and a less selfless desire to strike a blow against the newly founded People's Republic of China, insisted on voluntary repatriation for prisoners of war in Korea.[3] In other words, each individual soldier would be granted the right to refuse to return to his home country. The Communist Chinese government, mindful of the fact that the loyalty of many of its soldiers was questionable due to their previous service under the defeated and exiled Chinese Nationalist regime, vehemently resisted US demands despite the ruinous economic and personnel costs imposed by the continuing conflict. Both sides demonstrated their depth of ideological fervor by dragging out armistice negotiations for over two years until a final settlement was signed on 27 July 1953. In the end, the fears of the Chinese government were largely realized: over two-thirds (14,709) of the 21,629 Chinese POWs refused repatriation and defected to the Chinese Nationalist-controlled Taiwan.

The POWs' decisions to refuse or to accept repatriation were the final blows in a generation-long struggle between the Chinese Nationalist Party, the Kuomintang (KMT), and the Chinese Communist Party (CCP). From the mid-1920s, these two forces sought to construct a modern Chinese nation-state in either the fascist authoritarian or the communist mold, and each used the institutions at its disposal – the army, schools, and political party – to shape the loyalties of the Chinese people to help meet this goal. These institutions were designed to inculcate a strong sense of political identity, awareness and duty among the Chinese, especially among the

---

[3] Memo from Commander-in-Chief, United Nations Command to US Secretary of State, September 18, 1953. Office of Naval Intelligence, POW Desk, Operational Section, 1949-1954. RG 38/370. Memo from Vatcher to Greene dated 7 August 1952. Records of the General Headquarters, Far Eastern Command, United States Army. RG 554/290. US National Archives.

vast peasantry that had long been excluded from political affairs. For many years the Nationalists enjoyed the upper hand, dominating the heavily populated coastal areas, winning international recognition and the support of the United States, and pushing the Communists out to the western fringes of China. However, the Nationalists' inability to address the central social and economic challenge of land reform for the peasantry in conjunction with the invasion of China by the Japanese brought about a reversal of KMT fortunes despite efforts by the United States to bolster its ally (Spence 1990: 484-504). When civil war came to China in 1946, the Communists demonstrated their organizational superiority by consistently defeating larger Nationalist armies and incorporating their remnants into the burgeoning People's Liberation Army. In December 1949, Chiang Kai-shek fled with a fraction of his forces to Taiwan, where he set up a regime-in-exile. Thus, prior to being captured in Korea and having to make a difficult decision concerning their preferred repatriation location all Chinese soldiers in our sample had been exposed to a varying degree to the competing Nationalist and Communist nation-building projects. The question that drew us into this project was whether exposure to these competing state-building institutions left a lasting mark on the loyalties of the Chinese soldiers.

EVOLUTION OF THE RESEARCH PROJECT

A Dataset Based on US Archival Documents

The National Archives of the United States in Washington, DC, seemed like a logical place to look for answers, as we knew that the US Army compiled a large set of individual interrogation reports based on interviews with the Chinese POWs. The POWs were interviewed for the purpose of both obtaining actionable military intelligence and greater general knowledge about the political and social environment in the People's Republic of China (PRC). Interrogation reports were modeled after a template that had been refined during the Second World War and were therefore very

systematic. Specifically, each report contains detailed information on the interviewee's life history, including basic demographics – age, place of birth, education level, military rank and record of service – and chronology of life events including promotions in civilian and military occupations and membership in political organizations. The reports also contain a narrative section of varying length where soldiers recount their experiences in "New China" under the Communists. Information in the reports provides an empirical basis for a test of the institutional theory of political loyalty according to which men exposed to long-term socialization in Nationalist institutions (school, party, and army) should be more likely to refuse repatriation, and those who were socialized into Communist institutions should be more likely to accept repatriation.

On the basis of data contained in the interrogation reports and information concerning each soldier's repatriation decision we constructed a dataset containing 1,043 observations and over fifty explanatory variables and developed a model to test the various competing hypotheses (institutional, material interest, and familial attachment) that might explain the observed repatriation patterns. Given that institutional theories of loyalty formation are of primary interest to us, the nature of exposure to various institutions of the modernizing state (e.g. length of schooling, duration of service in the army, etc.) and one's status in those institutions (e.g. party member, Army rank, etc.) are the primary independent variables. The model also contains independent variables reflecting possible alternative explanations for the soldiers' repatriation behavior (age, family status, civilian occupation, etc.). The dependent variable is binary and is coded as 1 for those soldiers who repatriated and 0 for those who chose Taiwan over mainland China. In the interest of keeping the analytical spotlight firmly fixed on the central methodological point of this paper we will not go into an extensive discussion of our model or the various theories that inform it in this article – readers who are especially curious about the substance of this project are invited to consult our work elsewhere. Here we present only a truncated

and stylized description of our inputs and of the model estimation results.

Descriptive statistics reported in Table 1 paint the portrait of an 'average' soldier in our sample: he is young, modestly educated, and likely to have been either a peasant farmer or a student in civilian life. Importantly, four-fifths of the men in the sample served in the Chinese Nationalist Army (CNA) before joining the Chinese Communist Forces (CCF), and on average they spent more than a year longer serving in the CNA than in the CCF. Two percent of the men had served as senior officers in the Nationalist Army, and six percent held senior military or political positions with the Communists. Thirty-six percent of men in our sample repatriated; this figure is slightly higher than the 31% repatriation rate for the whole population of 21,629 Chinese POWs.[4]

The results of probit regression analysis largely confirm the institutional hypothesis, as we demonstrate in Table 2. Men who served with the Nationalists before joining the Communist forces are increasingly likely to refuse repatriation as the length of their period of service in the CNA increases. Whereas the average likelihood of repatriation for those in the sample with no service in the Nationalist Army is 57%, those with four to six years of service have an average repatriation rate of just 27%, and this rate falls to 20% for those who served in the Nationalist army for over six years (Table 3). By contrast, senior political officers in the CCF (i.e. those with considerable exposure to the Communist party and with a relatively lengthy record of service in the Communist forces) are on average 29 percentage-points more likely to

return to China than an average individual in the dataset. Likewise, CCF cultural instructors – mostly literate junior officers responsible for educating the troops – are 11 percentage-points more likely to repatriate than an average soldier. These coefficients are statistically significant in a model that includes all the standard controls for family status, class, age, and educational attainment.

At this stage we had a publishable paper on our hands. Our findings, based on a large-*n* dataset and econometric analysis, painted a theoretically plausible picture of individuals voluntary choosing one side over another based on the depth and length of their exposure to the fledgling institutions, particularly the army and the party, of the two competing state-building political machines. Our statistical analysis was methodologically sound and followed the standard best practice in the field. In fact, we were advised to go ahead and publish the paper as it stood. However, we began to have doubts about the validity of the data that we had collected up to that point. Though the interrogation reports were remarkably rich in both scale and scope, they covered the interviewees' life stories only up to the point of their capture on the battlefield and were therefore almost entirely silent on the prisoners' more recent experiences in the POW camps. Our next move was pivotal to the evolution of this research project: we decided to continue digging to reassure ourselves that the data that we compiled were in fact truly representative of reality.

At first we turned to other US archival sources. Even a careful examination of the narrative section of the interrogation reports which are not easily reducible to regression inputs – where the interviewees share their formative stories and talk about the Communist takeover of power in China – proved to be a highly useful exercise in testing for data reliability. From these narrative sections and from other Army and State Department documents we had a first tentative glimpse into complex reality that was somewhat at odds with our findings: it appeared that while many men were indeed highly ideological, a substantial subset of the prisoners seemed not to care about politics or to know much about it. Some of

---

[4] Generally, we believe that our dataset provides an appropriate baseline from which to extrapolate about the population at large. The main issue of concern is that US interrogators were more interested in interviewing better-educated men, because they believed that these men would provide intelligence of a higher quality (interview conducted by the authors with a former US army interrogator, 10/4/2008). Nevertheless, interrogators often found it difficult to establish right away which of the men were better educated, and the sample is sufficiently large for us not to be overly concerned with a substantial sampling bias.

TABLE 1. Descriptive Statistics

Descriptive Statistics	Mean	Standard Deviation	Benchmark Mean[†]	Avg. Prob. of Repatriation
% repatriating to mainland China	0.36	0.48	0.31[††]	0.36
Demographic				
Age	26.14	5.42	24.38	-
Years of education	5.13	3.50	2.41	-
*Economic** (% in each profession)*				
Farmer	0.29	0.45		0.37
Student	0.22	0.41		0.36
Tradesman	0.09	0.29		0.23
Laborer	0.09	0.29		0.39
Personal				
% Married (estimated by proxy)	0.23	0.42	0.28	0.38
% with no listed personal contacts on mainland	0.06	0.24		0.34
Military				
% Previously served in the Nationalist Army (CNA)	0.82	0.39	0.65	0.32
Average length of CNA service (yrs.)	3.44	3.57		-
% Served as a CNA senior officer	0.02	0.16		0.8
Average length of CCF service (yrs.)	2.10	1.22	1.22	-
% Served as a CCF senior officer	0.05	0.22		0.58
CCF political officer	0.01	0.12		0.71
CCF cultural instructor	0.13	0.34		0.52
CCF junior officer	0.21	0.41		0.38
CCF non-commissioned officer (NCO)	0.19	0.39		0.40
CCF enlisted man	0.40	0.49		0.24

Notes: N = 1019 (some observations missing)

[†]Benchmark figures from China General Survey (N=16,278) conducted by the US Army and State Department at the prison camps on Koje-do on March 12, 1952. The survey did not cover those confined in a work compound, a segregation compound for "Communist agitators", recently captured POWs still in the Pusan processing center, and hospitalized POWs [source: (Bradbury, Meyers, and Biderman 1968), Appendix A]

[††]Calculation from the final repatriation rosters complied by US Army (excludes deceased, escaped, and captured post-6/1/52 from the baseline)

TABLE 2. Probit Regression, Reporting Marginal Effects

DV = Repatriate	Marginal Effect	Standard Error
Key Variables		
Years in CNA	-0.022***	(0.006)
CNA Senior Officer	-0.213**	(0.084)
CNA Political Officer	-0.155**	(0.048)
Years in CCF	-0.008	(0.016)
CCF Senior Officer	0.192***	(0.089)
CCF Political Officer	0.293*	(0.150)
CCF Cultural Instructor	0.112**	0,058
Controls		
Landlord	0,091	(0.101)
Businessman	0,055	(0.085)
Farmer	0.098***	(0.039)
Married	0.008	(0.039)

* = p<0.05, ** = p<0.01,*** = p<0.001. N = 1,043

Pseudo-Rsq =

Note: Additional control measures, such as non-spousal family contacts or civilian occupations, beyond those displayed are not reported.

TABLE 3. Repatriation as a Function of Length of Service in the Nationalist Army

		Length of service in the Chinese Nationalist Army (years)					
	no CNA service	>0-2	2-4	4-6	6-8	8+	Total
% choosing repatriation	57%	43%	32%	27%	20%	22%	36%
Group size (# soldiers)	189	248	274	135	133	64	1,043

these men were bewildered youngsters or unsophisticated farmers with no interest in the ideological conflict between the Nationalists and the Communists, whereas others appeared to be opportunists who were willing to side with either power if offered sufficient incentives. Once again we started to wonder whether our statistical analysis did justice to the complexity of the situation – were most men driven in their choices by previous exposure to indoctrinating institutions of the rapidly modernizing state, or was some other latent variable responsible for the repatriation patterns that we observed?

Coercion in the camps – something that the interrogation reports did not cover – was one credible candidate for such a latent variable. Some US documents, especially those originating with the Prisoner of War Command and the Judge Advocate's General office, did indeed suggest that something strange was afoot in the camps. For instance, several important documents recorded a major upsurge in violence immediately preceding the repatriation screenings of October 1951 and April 1952. A number of prisoners were even murdered by their fellow inmates at that time, and investigations by US authorities failed to uncover the perpetrators of these crimes.[5] Although it was quite clear from these sources that violence was indeed being used instrumentally by certain factions within the camps, it was not at all obvious how much power these factions wielded within the prisoner compounds or whose interests they represented. The fact that two-thirds of the men refused repatriation suggests that

coercion was applied mostly by pro-Nationalist factions against Communist sympathizers, yet based on the evidence available we were in no position to preclude the existence of powerful pro-Communist factions. In short, we reached a dead-end with US archival sources; there was reason to suspect that we were seeing only part of the picture, but our suspicions were still very vague. It was time to diversify our data-gathering and data-analysis techniques.

Qualitative Interviews with Surviving POWs

We decided to interview some of the surviving protagonists next. Overall we interviewed 53 former POWs – 38 in Taiwan and 15 in China. Interviewees were difficult to locate because relatively few former POWs are still living, and an even smaller proportion is willing to speak openly about what was the most traumatic period of their life. Furthermore, particularly in China the state does not welcome a public discussion of any issues relating to the Korean War. Despite these difficulties, the interviews that we did secure were highly informative. Each interview lasted for approximately two or three hours; all interviews were open-ended, and we invited the veterans to share with us the memories of their upbringing, education, army service, experiences on the battlefield, life in the POW camps, and also the story of their post-war lives.

The central finding from our interview research is that many of the POWs were subject to coercion in the run up to repatriation screenings. Almost all the interviewees both in Taiwan and China related to us the story of what in many ways should be thought of as a reigniting of the

---

[5] Reports of Boards of Officers: POW Division. HQ, US Air Force Far East, Assistant Chief of Staff G-1: POW Incident Investigation Case Files. RG 554. US National Archives.

Chinese Civil War in the US-administered POW camps. UN troops captured substantial numbers of Chinese soldiers in January 1951, and by June of that year the informal leadership of most of the camp compounds housing the Chinese POWs was in Communist hands. Committed Nationalists among the prisoners set up a secret counter-organization, and the civil conflict that had been resolved in the Communists' favor in December 1949 in mainland China began to replay itself in the camp. During the fateful events of 1 October 1951 – when Nationalist and Communist sympathizers clashed in a massive fight over the raising of a KMT flag in the largest Chinese POW compound – Communist compound leaders and their most committed followers were made known to the US authorities and were rounded up and moved to a separate and much smaller segregated area within the camp. As a result, by the end of October 1951 informal leadership over the two largest compounds that together housed 14,000 of the total of 21,629 Chinese POWs passed to the Nationalists. Around the same time POWs were informed that they would be able to voluntarily express their preference for not returning to China.

The decisive repatriation screenings were held in April 1952. The interviewees on both sides of the conflict concurred that between October 1951 and April 1952 all 14,000 Chinese POWs under nominal Nationalist tutelage were subject to a sustained campaign of intimidation and violence designed to convince them to "voluntarily" choose Taiwan over China at the repatriation screenings. The process of coercion was multi-faceted and gradual. It would begin with leaders of smaller teams investigating their subordinates concerning their political loyalties – this practice was first introduced at the time when the Communists were still in the ascendance in the compounds. Those who agreed to express a preference for Taiwan were rewarded with better jobs and better food rations. The recalcitrants were forced to sing inflammatory songs and to donate blood that was used in place of ink for the signing of letters to the US and UN authorities denouncing repatriation. As the repatriation screenings drew nearer,

violence became more commonplace. Some of those who insisted on returning to China were tattooed with anti-Communist slogans (some of these slogans were in English lest doubt should arise in the minds of US screening personnel); others were beaten or publicly executed. One of the interviewees who "chose" Taiwan remarked thus on the activities of the Nationalist organization in the camp: "I left the Iron Curtain and ended up behind the Steel Curtain… They would kill you and not even bury you. They would burn you on a coal fire and throw your remains down a latrine."[6]

While many POWs felt coerced to demonstrate their allegiance to the KMT, coercion was certainly not unidirectional. Within their two smaller compounds containing about two thousand men in total the Communist leadership was pressuring soldiers to repatriate. Violence there never reached quite the same pitch as in the main compounds controlled by the Nationalists, because the Communist compounds by and large housed committed party members and officers of various ranks. The Communists also generally had an easier task because many of the apolitical men preferred to return to their homes by default. The remaining five thousand Chinese POWs, who were scattered across many installations managed by the US POW Command (at full capacity the camps were home to 102,000 men, mostly Koreans), were largely safe from political violence in their small teams of workers or convalescents.

All in all, interview research proved to be an absolutely essential complement to archival work. Through interviews we uncovered a vital latent variable – the fact that for many POWs decisions concerning repatriation were not voluntary, contrary to what we had previously assumed.[7] Also,

---

[6] Taiwan interview #32; 7/29/2009.

[7] Communist delegates at the armistice talks claimed all along that pro-Nationalist coercion was rampant in the POW camps and demanded rescreenings in the spring of 1953. It was then decided that all those rejecting repatriation should be rescreened by a neutral party, the Neutral Nations Repatriation Commission (NNRC). The NNRC was not up to the task of organizing the rescreenings because of organized protests by the Chinese POWs and

interviews made it apparent that the POWs were not an amorphous whole, but were rather split into four different types – two small pro-Communist and pro-Nationalist ideological minorities that were responsible for instigating political violence, a small but powerful group of violent criminals who were prepared to serve any master and proved to be a useful tool in the hands of the faction that enjoyed political preeminence, and the bulk of the soldiers who had either weak political preferences or no strongly articulated political views at all. In short, the interview findings challenged almost all of the assumptions that we formed in the first stage of data collection. It turned out that our statistical model was subject to an omitted variable bias, and more importantly that it was wrong to lump all the soldiers into one big data cauldron given that four distinct types of actors coexisted side by side in the camps.

Before going back to the drawing board to reconsider the overall research design we had to make sure that our interview findings were in fact reliable. To do this we turned to a set of unpublished interview transcripts that had been compiled by a senior Chinese army officer in the early 1980s (Jinfeng). The officer in question was deeply shaken by the fact that the former POWs who repatriated, whom he thought to be heroes for resisting pro-Nationalist coercion in the camps, were treated as spies and traitors by the Chinese government for some thirty years. On his retirement from the Army he decided to set the record straight and traveled across the country on his own savings to interview hundreds of former POWs. Notably, he also had access to transcripts of many of the screening interviews that each repatriate had to go through immediately on their return to China in the summer of 1953. In total, his work drew on information from 440 ex-

POWs. Sadly, for many years this officer was not able to share widely the transcripts that he prepared because of the political sensitivity of this topic. These transcripts are particularly interesting from our perspective, because he interviewed a number of the same men that we spoke with on our trip to China in 2009. It was gratifying to see that the story that his 440 interviewees confided to him was consistent with the information that we gathered on our trip.

Returning to the Drawing Board

Confident of the reliability of our interview data we could now return to the drawing board in order to refine our research design. We now knew that only a relatively small proportion of the POWs had strong political views – somewhere between 20-40% of the total; the remainder of the soldiers were largely apolitical. We also learnt that while most POWs felt that their repatriation decision was not entirely voluntary, nevertheless about 5,000 Chinese soldiers were in fact able to express their true preferences. It seemed logical therefore to focus our attention on individuals who resisted the dominant coercion effect – these are extreme cases of POWs who repatriated in the face of the dominant pro-Nationalist coercion or those who chose Taiwan even though they were housed in the Communist-controlled compounds. In addition, we also decided to salvage our original research design by limiting our regression analysis only to those individuals in our sample who were part of the substantial group of Chinese POWs not subject to any strong coercion effect – members of worker teams, convalescents, etc. In short, insights from interview work allowed us to reshape our quantitative strategy in such a way that the regression analyses we were now running had a considerably higher traction on reality.

WHY MULTIMETHOD RE-SEARCH PRODUCES BETTER DATA

We now use our specific account of voluntary repatriation to illustrate some general pathways by which a multimethod

---

also because of internal divisions within this organization, and so the mass rescreenings never took place. We omit a detailed discussion of this series of events as well as of an enormous amount of other fascinating material concerning life in the camps, as a fuller picture of the repatriation process is available in our other work.

research approach leads to improved data quality. The use of several different data collection techniques allows us to build a more balanced and complete dataset and affords a deeper insight into the way our data are structured. In this section, we focus on two particular pathways, arguing that a multimethod approach (1) increases the likelihood of finding and using sources on different sides of a given political cleavage to correct for political bias; and (2) produces a more complete dataset by overcoming the intrinsic limitations of a particular data-collection format.

In political science, much of the most interesting data tend to be generated by governments or political organizations of one sort or another, such as political parties, unions, or insurgent groups. Data generated by politicized organizations are often biased as political groups seek to exaggerate achievements and minimize failures. In the high-stakes arena of contentious politics – wars, protest movements, ethnic politics, labor struggles and so on – where the threat or use of violence casts a constant shadow, data frequently become a weapon to be manipulated for maximum political advantage. Notably, not only rebels and ideological rivals seek to propagate their version of social reality. Any political actor has an incentive to fudge data when political and economic stakes are high (cf. Greece, an OECD economy, and its chronic misreporting of national economic statistics in the run up to the 2010 international bailout). Although political actors do not always misrepresent reality, we must be leery of accepting data at their face value when they are generated by parties that have obvious incentives to promote their particular 'spin' on social reality.

Our project is a good example of how vigorously competing ideologies can act purposively to bias the reporting of an underlying process of interest. Each of the three political groups in the prison camps – the United States, the Chinese Nationalists and the Chinese Communists – had their own reasons for describing repatriation in a manner favorable to their political interests. With the barely-cooled embers of the brutal Chinese Civil War reigniting in the camps and the demonstrated willingness of the United States to risk the lives of its own soldiers by refusing to end hostilities in order to prevail in the ideological conflict over the principle of repatriation, the stakes were literally life-or-death and thus hardly conducive to an unbiased reporting of the facts. To navigate this or a similar type of contentious political situation one can either attempt to seek out neutral third-party sources or triangulate between readily-available, but highly politicized sources produced by actors that are directly party to the process under study. In our case, while third-party sources provided useful documentation (e.g., the reports of the International Red Cross), they could not deliver the granular, day-to-day information necessary to untangle the complex politics of repatriation. If we had restricted our analysis to our initial method, we would only have a secondhand representation of the views of the POWs, one filtered by the bureaucratic and political priorities of the United States military. Through interviews with the POWs we secured direct evidence from two other parties to this historical episode and thus were able to begin integrating the three political perspectives into a single narrative. As recounted in the previous section, analyzing all three points of view helped us not only identify clearly the latent variable of coercion, but also provided a roadmap for delineating its mechanism and spatial geography. Interviews with the repatriate POWs provided us with a clear map of the geography of coercion in the camps, validated by repetition from interview to interview; significantly, this spatial account was largely confirmed by the interviews with the nonrepatriate POWs. We were then able to explain the puzzling pattern of camp violence that showed up in the Judge Advocate General files – extreme violence in some areas, with relative quiet in others. Given that the Americans and the opposing groups of POWs each had their own political perspective, their agreement on the geography of coercion gave us a high degree of confidence that we had successfully balanced the competing political biases surrounding repatriation. Although triangulation between politicized sources can be self-defeating if all the sources are wildly misleading, the example that we provide here underlines the

practical utility of this exercise when it comes to generating specific connections between different sources. These connections give one a better understanding of whether a source should be included and, if so, help specify how sources might be fruitfully integrated. In our case, and more broadly, data triangulation of the kind described here allows for construction of a database that is politically balanced, because it transcends the biases of the parties that originally generated the data. Furthermore, such triangulation gives the researcher an insight into the way the data are structured. For instance, we learnt that what we originally believed to be a homogenous body of POWs actually consisted of several sets of very different actors.

Concern over the substantive bias of the sources that generate our data is not the only reason to use different methods for data collection in a single project. Every prominent data collection format – archival research, interviews, surveys, experiments, government statistics – comes with its own set of intrinsic limitations. Although these formats are all useful in their own right for generating an appropriate data foundation, they each have important limitations that complicate the goal of producing a robust and complete dataset. For instance, the evolving nature of many research projects prevents even the most farsighted researcher from including all relevant questions in her custom survey; similarly, the standardized nature of government-produced surveys limits the scope for addressing specific hypotheses of interest. Well-designed randomized experiments may generate ample data demonstrating that treatment A causes higher or lower levels of outcome Y, yet supporting data that illuminate the mechanism underlying *why* A changes Y are frequently in short supply. A straightforward way to overcome this problem is to combine data-collection formats through the use of multiple methods, particularly in those instances where the drawbacks of one format are compensated in complementary fashion by the advantages of another. Our project uses two methods in conjunction: archival research and interviews. By way of illustrating the broader methodological

point we discuss the general strengths and weaknesses of each approach and then analyze how their combination helps us to produce a more complete dataset and gives us an insight into how our data are structured.

As John Goldthorpe notes, archival evidence should be considered a "relic" because it consists of documents that provide only a residual record of previous political events or processes (1991: 211). For Goldthorpe, relics are characterized by two traits that cause problems for social-scientific analysis. First, their fixed nature necessarily limits any expansion of the dataset to address questions beyond the concerns that prompted the initial creation of the documents. Second, documents that end up in a given archive are often selected for preservation either by chance events or by an official redactor motivated by ideology, a desire to rationalize, or sheer bureaucratic inertia. Given these biases, Goldthorpe advises historically-minded social scientists to try to replicate or even replace archival data with contemporary fieldwork that can be more precisely tailored to theory and hypothesis testing. One problem with this prescription is that few organizations outside of the ambit of the state have the authority, access or capacity to amass the sheer scale of documentation required to track a complex political process such as, for example, voluntary repatriation. Government documentation is quite frequently the best data available for analyzing many political processes, and can neither be easily nor ethically replicated using contemporary evidence.

Oral interviews are similarly subject to their own set of analytical drawbacks and advantages. Free-flowing, open-ended interviews are by definition subjective exercises that are conditioned by political bias, emotions, cognitive limitations and memory lapses. At the aggregate level, interviews are subject to selection biases of their own because subjects may be unavailable due to mortality or for political reasons. Yet oral interviews remain a vital part of many researchers' toolkits, because they are useful for clarifying causal chains, identifying latent variables, and improving the

conceptualization and measurement of existing variables.

In our project, logistical, political, and psychological limitations prevented us from creating our own dataset based on a large and representative set of interviews of surviving POWs. We were therefore faced with the challenge of combining a large sample of US Army interrogation reports with a smaller set of contextual interviews. As we have discussed, the interrogation reports contain both highly systematic demographic and life-event data and lengthy narratives; their primary limitation is their silence to the events in the camp leading up to the repatriation decision. By contrast, the considerably less systematic interviews include material covered in the interrogation reports and also events in the prison camp. Combining the two allows us to achieve Goldthorpe's prescription of complementing archival research findings with contemporary fieldwork. Indeed, the fact that we possessed interrogation reports for some of the same men we interviewed allowed us to rigorously double-check the reliability of the interrogation reports, increasing our confidence in the integration of the two methods. The structure of coercion in the camps revealed by the interviews can now be used to partition our interrogation report dataset between the majority of men subject to a ferocious coercion effect in the main prison camp compounds and the substantial minority of about 5,000 men lucky enough to be located in less politicized areas of the camp. We are thus able to produce a more complete dataset that combines that systematic, large-scale nature of the interrogation reports with the contextual insights of nonrepresentative interviews. Combining the two methods helped ameliorate the standalone flaws of each in isolation. Although the challenges of incorporating other data formats may differ in their particulars, the important general point is that adding a data-collection format via an additional method will nearly always generate a set of productive new hypotheses, key pieces of evidence and changes in perspective that allow us to construct more complete datasets.

CONCLUSION

In this article, we have bracketed the philosophical war of method versus method and instead advocated the pragmatic use of a multimethod approach as a means for developing higher-quality data. The task of producing good data should be considered analytically prior to the subsequent methodological question of how to best analyze the data, and we argue that increased disciplinary attention to this task will produce increasing returns relative to incremental methodological innovations. To make our case, we have illustrated the way in which our own research project on the repatriation experiences of the Chinese POWs during the Korean War benefited from the integration of multiple methods. By mixing archival and interview methods based on different sources, we produced a more complete and politically balanced dataset with higher fidelity to the complex process of repatriation than could have been achieved by using either approach in isolation.

We conclude this article with a set of recommendations for scholars interested in using multiple methods to improve data quality. To set these recommendations in their proper context, however, it is necessary to revisit once again the philosophical divide between quantitative and qualitative methods. Although the focus on data quality sidesteps complex philosophical disputes over causality, we must still confront the issue of the relative contribution of each method to the overall layout of the dataset. The reader has undoubtedly noticed that our project uses intensive qualitative archival and field research ultimately in the service of expanding and refining a quantitative dataset. This should not be read as an endorsement of the view that qualitative data and insights should seek only to mimic or marginally improve a dominant underlying quantitative template. To the contrary, we believe that the nature of the research question should dictate whether a given multimethod project is based on a quantitative foundation with substantial qualitative refinements, a qualitative foundation with quantitative refinements, or

a creative synthesis characterized by equal qualitative-quantitative balance. In our case the research question – why, in the aggregate, two-thirds of Chinese POWs refused repatriation – dictates, in our view, a quantitative-dominant solution. The important and unique story of each POW is not diminished but rather magnified by recognition of the powerful shared forces of past military experience and coercion on repatriation. To be sure, each man grappled with these forces in his own way, yet all found themselves subject to them inside the camps.

Our project thus has the clearest lessons for scholars faced with the similar structural challenge of integrating qualitative data with a quantitative foundation. We recommend that scholars in this position make a concerted effort to make their data transparently accessible not only to those in their own corner of the research ecosystem, but also to those scholars who are interested in the same substantive issues from a different methodological perspective. Sustained interaction and integration with a qualitative (or different quantitative) perspective can vastly improve the way in which variables are conceptualized and measured, and often helps uncover crucial latent variables. Furthermore, a different methodological perspective on the same data can illuminate the underlying context of the dataset as a whole, properly situating it in political time or as part of a broader sequence of events. Conversely, scholars working from a qualitative perspective can similarly improve their data by stress-testing it against quantitative evidence. Qualitative scholars sometimes use quantitative data as the "conventional wisdom" that serves as the backdrop for interesting and counterintuitive claims. Yet quantitative research can provide additional benefits prior to the stage of making causal claims, helping the researcher determine what cases or observations lie within her scope of interest, which variables are deserving of further problematization, why some variables lend themselves better to standardization than others, and so on. This kind of qualitative-quantitative interaction certainly does not have to concede the dominance or superiority of a quantitative

dataset template in order to to benefit from quantitative insights.

Although this type of qualitative-quantitative collaboration is useful and productive, the increasing number of scholars who pragmatically combine methods for causal explanation offers an evolutionary opportunity for the discipline to expand conventional notions of what constitutes a dataset. Scholars willing to flexibly apply different methods to answer a substantive question are well placed to develop frameworks for integrating traditional data-set observations with the less-bounded observations from qualitative data, placing the latter on an equal footing with its quantitative kin. We encourage multimethod researchers to keep their powder dry on combustible issues such as the philosophical basis of mixing methods to assess causality and instead focus on redefining how and where qualitative and quantitative methods can be combined to improve overall data quality. Over time, investment in this research agenda will help to generate general patterns and guidelines for combining qualitative and quantitative forms of evidence. In addition to its direct benefits, this process provides a forum in which even those researchers who do not share a multimethod predilection can profitably share their methodological insights without engaging in sometimes counterproductive struggles over causality. Although a multi-generational effort to improve identification strategies has resulted in many advances, the discipline as a whole remains no closer to fundamental consensus on a number of key methodological issues. Our approach offers an alternative way forward, one based on improving data quality as a necessary prerequisite to causal identification. Given its potential returns – of which our project is one specific example – and platform for intradisciplinary cooperation, we hope that many social scientists will join us in pushing a new multimethod research agenda forward.

REFERENCES

Becker, Howard and Blanche Geer. 1957. "Participant Observation and Interviewing." *Human Organization* 16:28-32.

Bennett, Andrew and Bear Braumoeller. 2005. "Where the Model Frequently Meets the Road: Combining Statistical, Formal, and Case Study Methods. Working Paper."

Bradbury, William, Samuel Meyers, and Albert Biderman. 1968. "Mass Behavior in Battle and Captivity: The Communist Soldier in the Korean War." Chicago: University of Chicago Press.

Brady, Henry and David Collier. 2004. "Rethinking Social Inquiry: Diverse Tools, Shared Standards." Lanham: Rowman and Littlefield.

Brewer, John and Albert Hunter. 2005. "Foundations of Multimethod Research: Synthesizing Styles." Thousand Oaks, CA: Sage Publications.

Campbell, Donald and David Fiske. 1959. "Convergent and Discriminant Validation by the Multitrait-Multimethod Matrix." *Psychological Bulletin* 56:81-105.

Gage, Nathaniel. 1989. "The paradigm wars and their aftermath: A "historical" sketch of research and teaching since 1989." *Educational Researcher* 18:4-10.

George, Alexander and Andrew Bennett. 2004. *Case Studies and Theory Development in the Social Sciences.* Cambridge, MA: Harvard University Press.

Gerring, John. 2007. *Case Study Research: Principles and Practices.* New York: Cambridge University Press.

Goldthorpe, John. 1991. "The Uses of History in Sociology: Reflections on Some Recent Tendencies." *The British Journal of Sociology* 42:211-230.

Jinfeng, Wu. "A Survey Report on POWs in the Korean War. Unpublished Interview Transcripts. Translated by Cui Lin.". Beijing.

King, Gary, Robert Keohane, and Sidney Verba. 1994. *Designing Social Inquiry: Scientific Inference in Qualitative Research.* Princeton, NJ: Princeton University Press.

Lieberman, Evan. 2005. "Nested Analysis as a Mixed-Method Strategy for Comparative Research." *American Political Science Review* 99:435-452.

Lijphart, Arend. 1971. "Comparative Politics and the Comparative Method." *American Political Science Review* 65.

Mahoney, James. 2010. "After KKV: The New Methodology of Qualitative Research." *World Politics* 62:120-47.

Sartori, Giovanni. 1970. "Concept Misinformation in Comparative Politics." *American Political Science Review* 64.

Spence, Jonathan. 1990. *The Search for Modern China.* New York: W. W. Norton & Company.

Symposium. 2007. "Multi-Method Work, Dispatches from the Front Lines." *Qualitative Methods: Newsletter of the American Political Science Association Organized Section for Qualitative and Multi-Method Research* 5:9-28.

Tashakkori, Abbas and Charles Teddlie. 2002. "Handbook of Mixed Methods in Social and Behavioral Research." Thousand Oaks, CA: Sage Publications.

Trow, Martin. 1957. "Comment on participant observation and interviewing: A comparison." *Human Organization* 16:33-35.

US Army primary documents

Documents from the National Archives in Washington, DC (record group number in parentheses)

- Far Eastern Command/United Nations Command (RG 554)
- Records of the Army Staff (RG 319)
- Office of Naval Intelligence (RG 38)

Center for Advanced Study in the Social Sciences (http://www.march.es/ceacs/ingles/ceacs.asp)
WORKING PAPERS

Series Editor: Andrew Richards

Most recent titles:

Working Paper	Author	Title
2010/249	Kselman, D. M.	*Electoral Institutions, Legislative Accountability, and Political Corruption.*
2010/250	Pop-Eleches, G. and Tucker, J. A.	*After the Party: Legacies and Left-Right Distinctions in Post-Communist Countries.*
2010/251	Urquizu, I.	*Moderate Voters and Internal Party Factionalism.*
2010/252	Kuo, A. and Margalit, Y.	*Measuring Individual Identity: Experimental Evidence.*
2010/253	Kuo, A.	*Explaining Employer Coordination: Evidence from the US Inter-War Period.*
2011/254	Falcó-Gimeno, A.	*Portfolio Allocation and Time out of Office in Coalition Governments.*
2011/255	Gorodzeisky, A. and Semyonov, M.	*Occupational Incorporation of Immigrants in Western European Countries.*
2011/256	Brader, T. A. and Tucker, A.	*Follow the Leader: Party Cues, Policy Opinion, and the Power of Partisanship in Three Multiparty Systems.*
2011/257	De la Calle, L. and Roussias, N.	*Independents and Vote Choice: Spatial or Performance Voting?*
2011/258	Lapuente, V.	*Trade-Offs in Corporate Governance.*
2011/259	Lago, I.	*Why (Not) National Party Systems?*
2011/260	Burgoon, B.	*Immigration, Integration and Support for Redistribution in Europe.*
2011/261	Kelly, J., Hamann, K., and Johnston, A.	*Unions against Governments: Explaining General Strikes in Western Europe, 1980-2006.*
2011/262	Riera, P.	*Electoral Coordination in Mixed-Member Systems: Does the Level of Democratic Consolidation Matter?*
2011/263	Aguilar, P.	*Authoritarian Repression, Judicial System and Transitional Justice: The Spanish Case in Comparative Perspective.*
2011/264	Cebolla, H. and Finotelli, C.	*Integration beyond Models: An Empirical Outlook to the Impact of Integration Models.*
2011/265	Rivero, G.	*Integrality and Separability in Multidimensional Voting Models: Ideology and Nationalism in Spanish Regional Elections.*
2011/266	Hierro, M. J.	*Parents and School: Agents of `National´ Socialization?*
2012/267	Fernández-Albertos, J., Kuo, A., and Balcells, L.	*Economic Crisis, Globalization, and Partisan Bias: Evidence from Spain.*