Unifying the Study of Asymmetric Hypotheses

Andrew S. Rosenberg, Austin J. Knuppe, and Bear F. Braumoeller*

January 11, 2017

ABSTRACT

This article presents a conceptual clarification of asymmetric hypotheses and a discussion of methodologies available to test them. Despite the existence of a litany of theories that posit asymmetric hypotheses, most empirical studies fail to capture their core insight: boundaries separating zones of data from areas that lack data are substantively interesting. We discuss existing set-theoretic and large-N approaches to the study of asymmetric hypotheses, introduce new ones from the literatures on stochastic frontier and data envelopment analysis, evaluate their relative merits, and give three examples of how asymmetric hypotheses can be studied with this suite of tools.

Word Count: 8,989

*Department of Political Science, The Ohio State University, Columbus. The authors are grateful to Jose Fortou, Christopher Gelpi, Oul Han, Marcus Kurtz, Adam Lauretig, Richard McAlexander, William Minozzi, Anna Meyerrose, Jason Morgan, Adam Ramey, Gregory Smith, Avery White, and the participants in our panel at the 2015 American Political Science Association conference for feedback and comments on previous drafts. Replication materials are available at doi:10.7910/DVN/9NXGHP. These techniques are implemented in the open-source software `asymmetric`, which can be downloaded at https://github.com/asrosenberg/asymmetric-package. Further examples and resources can be found at https://www.asrosenberg.com/asymmetric-package/.

INTRODUCTION

Political science is making welcome progress out of what Morgan and Winship (2014, 10–13) call the "Age of Regression"—the dark period between quasi-experiments and causal inference in which nearly every claim was tested with a linear model and concerns about confounders were addressed with the addition of yet another control variable. Today's scholars have available to them a much wider array of statistical methods that conform far more closely to the logic of their theories. This is exceptionally good news for political science.

That said, there are still a handful of areas in which statistical inference is guided more by residual linear-additive thinking than by the logic of the problem at hand (Abbott 1988). Foremost among these are tests of *asymmetric hypotheses*—situations in which the hypothesized role of X is to establish a limit on the variation of Y (Lieberson 1985, ch. 4). Examples of these hypotheses can be found in every subfield of political science. The literature on the democratic peace (Maoz and Russett 1993; Russett 1994; Gartzke 1998; Bueno de Mesquita et al. 1999) is an obvious one: the presence of dyadic democracy places an upper limit on the hostility of interstate relations.[1] The distinguishing feature of these hypotheses is the conspicuous absence of observations from a region of the data space.

In this article, we provide a reevaluation of the methodology of asymmetric hypotheses. We argue that, because the absence of observations is the phenomenon of interest, data boundaries are more theoretically interesting than the central tendency of the data or a hypothesis of no effect. Therefore, we present a suite of tools useful for data exploration, rather than hypothesis testing. These tools allow one to discover boundaries that separate the region in which observations are present from the region in which observations are mostly or entirely absent and to gauge the uncertainty around those boundaries.

[1] A more traditional description of this relationship would be that dyadic democracy is sufficient for peace, or conversely that the absence of democracy in at least one country in a pair is a necessary condition for war between them. While not unreasonable, this formulation is deeply unfortunate. The words "necessary" and "sufficient" conjure up the bogeyman of deterministic causation—an ontological premise that, in our experience, is immediately rejected by every scholar with a background in statistical methods and causal inference techniques (or, to avoid argument, by almost every such scholar). As a result, asymmetric hypotheses receive little attention from the quantitative methods community and have mostly been shoehorned into existing methodologies, such as tests of heteroskedasticity and GLMs with multiplicative interaction terms (Clark, Gilligan and Golder 2006). While we find the discussion of deterministic relationships to be something of a red herring (Braumoeller and Goertz 2000; Dion 1998; Ragin 2000), we do avoid a discussion of causation because the techniques we highlight below neither ensure nor preclude a causal relationship.

One can use these tools to test hypotheses or to preprocess and explore data as a complement to further analysis and theory-refinement (Seawright 2016).

We should emphasize the fact that, although this is a methodological article, its main contribution is conceptual. The methods used to estimate boundaries, while helpful in that capacity, are most useful as concrete illustrations of a fruitful perspective on social phenomena, the asymmetric relationship, that cannot be fully conceived from within the confines of the linear model (or, for that matter, the average treatment effect). While the methods will certainly improve empirical research, our main hope is that widespread comprehension of asymmetric hypotheses will fuel more nuanced theorizing and motivate further methodological progress.

This article proceeds in five parts. In the next section, we motivate the discussion of asymmetric hypotheses. Following that, we evaluate the current state of the art for testing asymmetric hypotheses: multiplicative interaction terms and set-theoretic approaches. Third, we describe three lesser-known econometric techniques that incorporate the best features of these methods while avoiding their drawbacks. We then present three examples of how asymmetric hypotheses can be studied with this suite of tools. We conclude by discussing different techniques that can be used to estimate this boundary, and we illustrate their strengths and weaknesses in the context of three different asymmetric hypotheses.

Motivating the Study of Asymmetric Hypotheses

The robust relationship between district magnitude and electoral system size is one of the most well known and influential findings in all of social science. It is also an outstanding example of an asymmetric hypothesis. In Duverger (1963)'s canonical presentation, the author argues that single-member district plurality (SMDP) electoral systems are sufficient to produce a two-party system—or equivalently, that a non-SMDP electoral system is necessary to produce a multiparty system.

The hallmark of asymmetric relationships is quite dramatic: cells of a table, or regions of a scatterplot, that are conspicuously devoid of observations. Importantly, when X is continuous, asymmetric relationships typically produce a "floor" or a "ceiling" below or above which observations are rarely if

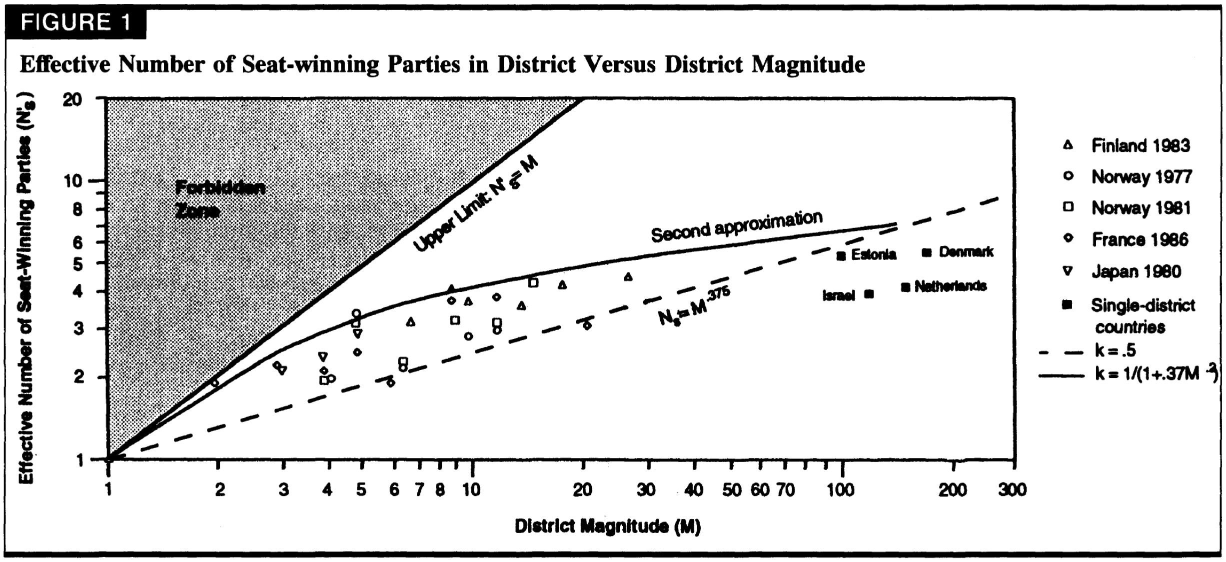

Figure 1: A test of Duverger's Law: the relationship between district magnitude and the effective number of parties (Taagepera and Shugart 1993).

ever found (Goertz, Hak and Dul 2013). Noting the absence of observations in a scatterplot redirects the analyst's attention toward an important feature of the data that often goes overlooked but is almost always of great theoretical interest: the boundary between zones of data and no-data.

Like many before them, Taagepera and Shugart (1993) and Li and Shugart (2016) test the empirical implications of this relationship. While Taagepera and Shugart (1993) discover that, consistent with the original hypothesis, there is an upper limit on the effective number of seat-winning parties in districts of certain magnitude, Li and Shugart (2016) find that this is upper limit is better predicted by alternative institutional variables. These results serve as an excellent motivating example because both studies are explicitly interested in estimating the upper limit of observations that can be seen in in Figure 1. In other words, neither Duverger nor the authors is especially interested in the central tendency of the data or the residual distance between observations and the regression line. Rather, their focus is on the region of the data space—the ominous "Forbidden Zone" in Figure 1—in which, according to Duverger, few if any observations should be found. They seek to estimate the placement and characteristics of a line that describes the upper limit of the set of observed cases, rather than the central tendency of the data.

The ubiquity of these "Forbidden Zones" has motivated scholars to turn toward the study of "asymmetric" hypotheses (Clark, Gilligan and Golder 2006). An asymmetric relationship is one in which the primary theorized role of X is not to increase or decrease the average value of Y—though it may do so incidentally—but rather to establish a limit on its variation. It is the maximum (or minimum) value of Y, rather than its average value, that varies as a function of X. Returning to the Duverger example in Figure 1, the value of X (district magnitude) has an impact on Y (the effective number of seat-winning parties) not by changing its average value but rather by providing an upper limit to its variation: in large districts, there may be many parties or there may be few, but as district size decreases so too does the maximum number of effective seat-winning parties.

Exploring data boundaries has substantial implications for how asymmetric hypotheses are studied. While symmetric hypotheses lend themselves to estimating the central tendency (i.e., the conditional mean) of the data, asymmetric hypotheses are concerned with its limit. Put another way, investigating data boundaries is not reducible to merely comparing E[$Y|X$ = 0] and E[$Y|X$ = 1] (or E[$Y|X$ = n]) because these quantities say nothing about the largest or smallest possible value of Y we can expect for given values of X. For example, a theory about how, *on average*, an increase in district magnitude leads to an increase in party system size misses an enormously important fact about the world if the absence of such magnitude *limits* the formation of multiparty systems. If scholars focus on these latter sorts of relationships, then they will be able to leverage a rich and diverse set of tools to test their more nuanced theories in a principled way.

Asymmetric hypotheses, and the boundaries they produce, are of interest to virtually every subfield of political science. In Table 1 we provide several illustrative examples of prominent studies of asymmetric hypotheses across the discipline. In each case, we note when outcomes are either predicted well or predicted poorly. Our goal is to show that, regardless of one's methodological persuasion, there is a common logic that underpins asymmetric hypotheses that all political scientists share. The exciting prospect of a renewed, unified focus on asymmetric hypotheses is that many scholars seem to think in these terms, and there are a variety of methods, quantitative and qualitative, for studying them.

In the next section, we review commonly used methods for studying asymmetric hypotheses. We first discuss the pros and cons of set-theoretic approaches. Then we show, in agreement with Clark, Gilligan and Golder (2006), that there is nothing about large-N statistical techniques that precludes them from being used to study asymmetric hypotheses. However, we argue that multiplicative interaction terms are ill-suited to the task because they are designed to estimate the central tendency of the data rather than its upper or lower bound.

ASYMMETRIC HYPOTHESES: THE STATE OF THE ART

Two methodologies define the current state of the art when it comes to testing asymmetric hypotheses: Qualitative Comparative Analysis (QCA) and multiplicative interaction terms. Below, we outline each and highlight their strengths and weaknesses before moving on to a discussion of methods that combine the strongest points of each.

Set-Theoretic Approaches: Varieties of Qualitative Comparative Analysis (QCA)

Set-theoretic techniques, such as csQCA and fsQCA (Ragin 1987; 2000; 2008), have come under fire in recent issues of *Political Analysis* (Hug 2013; Braumoeller 2015; Krogslund, Choi and Poertner 2015). Yet their proponents should be lauded for their focus on the key quantity of interest in asymmetric hypotheses: boundaries between zones of data and no-data.

Set-theoretic approaches focus on the absence of data by modeling which of the logically-possible combinations of qualitative conditions preclude certain outcomes, as opposed to quantitative techniques which measure the extent to which independent variables covary (Thiem, Baumgartner and Bol 2016). Even if two variables covary, that does not mean that in *combination* they are either necessary or sufficient for an outcome. This observation leads Ragin to argue that "…the variable-oriented strategy is incapacitated by complex, conjunctural causal arguments" (Ragin 1987, 69). In short, QCA advocates argue that asymmetric hypotheses can be best understood using set-theoretic methods rather than regression-based techniques.

While set-theoretic techniques have gained popularity since Ragin (1987)'s early work, critics argue that the list of assumptions for QCA is long and quite restrictive and that the technique tends to produce unreliable results. Seawright (2005) points out that QCA can only uncover relationships that are either additive linear or multiplicative linear. Hug (2013) shows that QCA can produce misleading inferences if measurement error is not taken into account. Even more problematic, critics have shown that QCA tends to produce false positive results (Braumoeller 2015). Finally, Krogslund, Choi and Poertner (2015) cast further doubt on QCA by showing that a random variable added to a model can often be identified as a critical component of a causal relationship.

The main issue with fsQCA for our purposes is that the boundary of the data-free zone is imposed by the researcher and not empirically estimated. While researchers can calibrate the values of X and Y based on substantive knowledge of cases, the boundary between the data and the region in which few or no observations should be found is the line $Y = X$. Yet, as far as we know, no asymmetric theory is specific enough to justify the *a priori* assumption that that boundary lies at $Y = X$ (as opposed to, say, $Y = \frac{X}{2}$ or $Y = X^2$). The boundary should be estimated rather than assumed. These issues and assumptions make set-theoretic techniques less than ideal for studying asymmetric hypotheses.

Interaction Terms

In the quantitative tradition, Clark, Gilligan and Golder (2006) (hereinafter, CGG) offer the most comprehensive description of asymmetric hypotheses and justification for why they are vital to social science. They claim that previous attempts to study asymmetry suffer from two different pathologies: the techniques either take a deterministic view of mechanisms in the social world (Ragin 1987; 2000; George and Bennett 2005), or require outcomes to be monocausal (Braumoeller and Goertz 2000). CGG argue that a more appropriate technique for testing asymmetry would take into account the stochastic nature of the social world. They argue that the use of interaction terms in the standard linear model below is capable of capturing these effects:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_1 X_2 + \epsilon$$

Although CGG correctly diagnosed the problem, there is a serious flaw in their proposed solution: it

Citation	Predicted Well	Predicted Poorly
Moore (1966)	Development of Western-style democracy (will not happen) absent a bourgeois revolution from above	Development of democracy when there is a bourgeois revolution
Bueno de Mesquita (1983, 129)	The initiation of serious international disputes (will not happen) if leaders do not have positive expected utility	The initiation of serious international disputes if leaders do have positive expected utility (may or may not go to war)
Krasner (1983, 189)	The formation of security regimes (will not happen) when major stakeholders do not prefer the status quo	The formation of regimes (may or may not occur) when preferences of major actors prefer the status quo
Lijphart (1990, 488)	Adjusted district magnitude (is larger) in an L.R.-Hare and Pure S-L System	Adjusted district magnitude in an L.R. Droop, Mod. S-L, or d'Hondt system
Russett et al. (1995, 167)	War (will not occur) when there no irrationality or misperception	War when there is either irrationality or misperception
Mercer (1996, 6)	An opponent's reputation (will not form) when another actor makes a situational attribution	An opponent's reputation (may form) when an actor makes dispositional attributions
Gartzke (1998, 9)	Crisis escalation (none) when neither actor demonstrates either opportunity and willingness to fight	Crisis escalation (may occur) when both actors demonstrate both opportunity and willingness
Lupia and McCubbins (1998, 69)	Enlightenment (none) if speaker is not persuasive, not uniquely possessed of knowledge that the principal needs, or not compelled to reveal what he knows	Enlightenment if all three conditions are met (could occur, but might not)
Tsebelis (1999, 596)	Number of significant laws (few if any) when government has large range	Number when government has small range (could be many or few)
Schimmelfennig (2001, 69)	EU enlargement (none or very little) when there is not a community of shared political values and norms with outside states	EU enlargement when there is such a community of shared political values and norms
Kaufman, Little and Wohlforth (2007, 15)	The emergence of hegemony (will not happen) when there is not at least one effective empire	The emergence of hegemony when there is at least one effective empire
Wimmer (2014, 8)	Democratization (none) with high levels of exclusion	Democratization (may occur) with low levels of exclusion
Talmadge (2015, 229)	Battlefield effectiveness (little) when there is no cohesion within units	Effectiveness (militaries may or may not have it) when there is cohesion in a military organization

Table 1: Examples of asymmetric hypotheses from the political science literature. Column 2 shows the range of input that predicts the outcome well, whereas column 3 describes the range that does not.

fails to model the appropriate attributes of the data.[2] What is being modeled in an OLS regression is the central tendency of the data—the average value of the dependent variable, conditional on covariates. What distinguishes asymmetric from symmetric hypotheses is the fact that the former produce distinct zones of data and no-data. Accordingly, the quantity of interest in the case of an asymmetric hypothesis is the upper or lower *boundary* of the data. As we demonstrate below, CGG's own example of the determinants of party system size is better conceived of as a data-boundary question rather than a central-tendency question.

HAVING THE BEST OF BOTH WORLDS

We have just argued that set-theoretic methods capture the concept of asymmetric hypotheses very well but are fraught with inferential peril, while multiplicative interaction terms model the central tendency of the data rather than their limits. For those reasons, we cannot recommend the use of either for evaluating asymmetric hypotheses.

Below, we describe three methods that combine the best features of these two techniques—a fundamental recognition that the boundary between the region in which data are present and the region in which they are nearly or entirely absent is the quantity of interest and the ability to estimate rather than assume that boundary.

Quantile Regression

Goertz, Hak and Dul (2013) present quantile regression (QR) as a regression-based approach for modeling data thresholds with error. They suggest that asymmetric hypotheses require that analysts turn their attention away from conditional means and toward quantiles, which represent "ceilings" and "floors" delineating zones of data.

[2] While we claim that CGG do not model the correct attributes of the data, Thiem, Baumgartner and Bol (2016, 759) alternatively contend that interpretations of regression coefficients run counter to the logic of "Boolean-algebraic hypotheses."

As the authors explain, "The key principle to note is that ceiling and floor hypotheses involve a fundamentally different orientation to hypotheses and data analysis: Ceiling and floor hypotheses are about drawing boundary lines between zones of data and zones of no data; they are not about drawing lines through the middle of data" (Goertz, Hak and Dul 2013, 20). Ceiling and floor hypotheses are typically asymmetric hypotheses: they imply that the range of possible observed values change with different values of independent variables (a rich state has a lower limit on its level of democracy, for example (Goertz, Hak and Dul 2013, 20)).

QR works well for asymmetric hypotheses because it does not model the relationship between X and the mean of Y. Rather, it models the relationship between X and some pre-specified boundary on the value of Y, often the 95th (upper boundary) or 5th (lower boundary) quantile. QR is also exceptionally robust to different distributions of data, a feature that allows it to produce reasonable estimates when other techniques fail.[3] In addition, the regression line for a particular quantile demonstrates an important point that CGG also discuss: the treatment of counterexamples. In substantive terms, a "ceiling" regression line through the .95 quantile means that five out of every 100 observations are treated as counterexamples, or cases which stand outside of the asymmetric hypothesis derived from the theory, *a priori* (Goertz, Hak and Dul 2013, 29).

Thus, QR retains the intuition of CGG that it is necessary to model asymmetry with an understanding of the stochasticity of the social world. Yet, it is an improvement because QR explicitly focuses on estimating the object of interest for asymmetric hypotheses—data boundaries—and provides an improved approach to the question of error and counterexamples.

While QR shows that it is possible to use regression-based techniques to test asymmetric hypotheses, it is vulnerable to one criticism: while the threshold is estimated, the exact quantile to use to estimate it is assumed *a priori* rather than derived from the characteristics of the data. The scholar is given great latitude in determining what amount of error "seems right." While Goertz, Hak and Dul (2013) provide a procedure for determining which quantile should be estimated (the Optimum

[3] There is a vast quantile regression literature, both frequentist and Bayesian, that demonstrates the flexibility and robustness of the technique (see, Yu and Moyeed (2001), Koenker (2005, ch. 2), Angrist and Pischke (2008, ch. 7)). However, in all cases, the analyst is left to specify the quantile through which the regression line is drawn.

Boundary Line), it is cumbersome, particularly once extended beyond bivariate settings, and cannot be directly compared across studies. It is telling that even after discussing the OBL technique, they admit that despite finding it useful for guiding which line should be used, "we eventually want to make some comparisons across studies. One way to do this is to take a fixed standard...an obvious choice is the .95 quantile regression" (Goertz, Hak and Dul 2013, 29). So, in practice, the analyst is often left to convention when deciding where to place the boundary.

Stochastic Frontier Analysis

Stochastic Frontier Analysis (SFA) is a modeling technique from the field of economics that is designed to estimate cost and production frontiers rather than expected values. The canonical example takes the form $y_i = f(X_i; \hat{\beta}, \xi_i)$, where f is the production frontier, X_i are the explanatory variables, $\hat{\beta}$ is the vector of parameters of the production function to be estimated, and ξ_i, is a compound error term that adds to the normal, Gaussian error (v_i) of the standard regression model a representation of inefficiency for each observation ($-u_i$) (Aigner, Lovell and Schmidt 1977; Meeusen and Van Den Broeck 1977). This error term allows SFA to take the burden of quantile or counterexample specification out of the analyst's hand and put it into a familiar estimation routine that looks much like the standard linear model.

Most economists who use SFA take the production frontier to be in the Cobb-Douglas form, where the output of firms are a multiplicative function of land, labor, and capital. Although the Cobb-Douglas is illustrative because the frontier represents the theoretical maximum level of production for a firm, it is not ideal for most political science questions because, by definition, the absence of one of the factors of production leads to zero total production. The general SFA model is flexible enough to handle different specifications of f, but the shape of the boundary is most often assumed to be linear (Simar and Zelenyuk 2011, 3). Importantly, it models outcomes that look much like an asymmetric hypothesis with heteroskedastic data: firms can theoretically produce up to the frontier, so the zone above the frontier should contain few or no observations.

However, data boundaries are not absolute. Some observations are above or below frontiers due to inherent noise around the estimated boundary, or conceptual imprecision that results in measurement error. Herein lies another aspect of the value-added of SFA: while fuzzy-set techniques model conceptual imprecision, the handling of measurement error is left to the analyst; however, in SFA the routine *estimates* the location of the line and, by extension, how many observations are allowed to fall on the other side due to error that is familiar from other regression models. This corresponds to the first component of the compound error term in SFA models, commonly denoted as v_i. In the canonical example, land, labor, and capital are the three inputs posited to produce output, but there is some inherent randomness around the frontier (i.e., "noise"), that corresponds to the familiar ϵ term from linear models.

In addition to the inherent stochasticity around the frontier, there is a second source of randomness that enters into the model. Even if a firm has an ideal mix of the three inputs, it could still under-produce due to inefficiency from exogenous shocks, bad luck or any number of other factors (as well as over-produce for converse reasons). SFA is ideal for estimating data boundaries because its estimates include both a standard Gaussian error and a representation of how much or little a particular observation deviates from the production frontier. This leaves the final model to take the form of $Y_i = \beta_0 + \beta' X_i + v_i - u_i$ (Greene 2008; Simar and Zelenyuk 2011).

In this manner, SFA deals with the thorny question of counterexamples and the complex nature of the social world. Just as firms may out-produce what a production frontier would predict, states can also have greater levels of democracy than their level of GDP would indicate. Accordingly, SFA recovers the same insights that QCA does by using statistical inference rather than assumptions.

SFA is not without its own assumptions, of course. The main assumption of interest involves the compound error term. The distributional assumption of the error around the frontier, v, is familiar, as it is assumed to follow a Gaussian distribution with zero mean. A variety of distributions could be used to model the second, non-positive inefficiency error that accounts for the amount each observation fails to reach the frontier: typically, u can be distributed half-Normal, exponential or gamma, depending on the data (Greene 2008). The variance of u can also be modeled as a function of X. While this flexibility is welcome, it is also limiting, in that inefficiency terms that do not conform to these distributions can produce irregular results. If the distribution of the errors is too left-skewed, for example, the procedure may not be able to successfully estimate the inefficiency.

In sum, SFA is preferable to fsQCA and interaction terms because it allows analysts to estimate threshold surfaces with a confidence interval that provides not only a measure of uncertainty around the exact location of the threshold but also around the estimate of the inefficiency of each observation relative to the boundary via the compound error term. However, SFA comes with costs as well. It requires rigid functional form assumptions for the boundary and the error terms, and it can be fragile if those assumptions are not met.

Nonparametric Frontier Models

Another option for estimating data boundaries are "nonparametric frontier models" (NFM) such as data envelopment analysis (DEA). These models also attempt to estimate a frontier from observed data, but because these models are nonparametric there are no rigid functional form assumptions. Unlike SFA, no parametric restrictions (linear or otherwise) are made about the distribution of inputs on the boundary. We discuss classical DEA prior to presenting more recent developments.

DEA is a procedure invented by Farrell (1957) for use in operations research and organization science (Seiford 1996). It is primarily used to model productive efficiency of decision making units (e.g., the U.S. Post Office) when a production process involves a structure of multiple inputs and outputs (e.g., delivery of the daily mail). Here, the data boundary is the line which defines the maximum combination of outputs that can be produced for a given set of inputs. Observations that lie below the frontier reflect some degree of inefficiency in the production process.[4]

Formally, the data boundary (Ψ) can be expressed as

[4]Producers can either be "technically" inefficient (Koopmans 1951; Debreu 1951, 60), which means that they obtain less than the maximum output from the available array of inputs, or "allocatively" inefficient, because they fail to purchase the best set of inputs given their price and productive potential.

$$\Psi = \{(x,y) \in \Psi | (\gamma^{-1}x, \gamma y) \notin \Psi \text{ for any } \gamma > 1\},$$

where (x,y) is any ordered pair that falls on the boundary and γ is a constant. So, for any ordered pair on the boundary, as well as any constant $\gamma > 1$, no other ordered pair that multiplies x by γ^{-1} and y by γ can also be on the boundary.

Nonparametric boundary estimation requires three minimal assumptions about the boundary. First, one assumes that inputs and outputs are freely disposable, which simply means that increasing inputs in isolation will not lead to a decrease in outputs. Further, any level of input that produces a given level of output can also produce less of it, respectively. Second, the "no free lunch" assumption states that production is impossible without the use of some positive inputs. Third, the boundary set Ψ is assumed to be convex, which follows from the convexity of preferences in economics. In other words, for any two bundles of an input and output that are each on the frontier, a weighted average of the two bundles is also on the frontier.[5]

The original DEA estimator uses all three assumptions to envelop data without imposing a parametric structure on the distribution of (X, Y) on Ψ. However, other estimators exist that do not require all three assumptions. The Free Disposal Hull (FDH) estimator, for example, extends DEA by maintaining the free disposability assumption while relaxing convexity (Deprins, Simar and Tulkens 1984). The convexity assumption requires the boundary to make predictions about all possible values of X, not just those that are observed. When convexity is relaxed, predictions are only made about cases that are *observed*, leading to more persuasive estimates (Ray, Kumbhakar and Dua 2015). In FDH, the data boundary is estimated using the lowest "staircase" monotone curve covering all the data points. As a result of this "staircase" approach, FDH algorithms are easy to compute.

While DEA and FDH have the benefit of not assuming a particular functional form for the frontier (Simar and Wilson 2013), the assumption that all observations are attainable means that no noise is assumed in the data generating process. In other words, the distance between an observation and the

[5] If the ordered pairs $(x_1, y_1), (x_2, y_2)$ are on the boundary (Ψ), then for all $\alpha \in [0, 1]$, $(x, y) = \alpha(x_1, y_1) + (1 - \alpha)(x_2, y_2) \in \Psi$.

frontier is due only to inefficiency, not measurement error, and no observations are permitted to fall outside of the frontier. In addition, DEA and FDH estimate frontiers that fluctuate widely if there are outliers or other extreme values in the data. Relatedly, these techniques do not estimate smooth boundaries. Rather, DEA and FDH produce boundaries with many non-smooth points where "the estimated curve repeatedly stays flat and jumps, and the location of jumps [depends on particular X values and] has no appropriate interpretation" (Noh 2014, 503). Recently, DEA and FDH have been extended to remedy these discontinuities by accounting for measurement error, and these innovations provide a more flexible framework for the study of asymmetric hypotheses.

Accounting for Measurement Error. Recent innovations in NFMs allow one to retain the computational ease of DEA and FDH while simultaneously accounting for measurement error. While a fully nonparametric model with both noise and inefficiency is not identifiable (Hall and Simar 2002, 523),[6] NFMs can be amended to account for unusual observations. The beauty of these extensions is that they are incredibly robust and widely prolific—one can choose no fewer than ten different modifications of data envelopment techniques.[7] These extensions outperform DEA and FDH in the presence of noise (Simar 2007), and are less sensitive to extreme values in the data (Simar and Zelenyuk 2011; Daouia, Florens and Simar 2012).[8] In addition, each of these extensions has reasonable criteria for drawing boundaries that allow for measurement error: they build error into the techniques by considering all points within some bandwidth around the edge of the data to be drawn from a probability distribution. Finally, these techniques have a unique and robust way of dealing with inefficiency—they do

[6]The problem of two error terms does not allow for a unique solution to the problem of finding a θ that marks the location of the frontier, without parametric constraints see Hall and Simar (2002, 523) for technical details. Current research seeks to overcome this computational issue by imposing minimal structure on the model. See Simar and Wilson (2013, 321–326) for details.

[7]They are: Data Envelopment Analysis (Farrell 1957; Charnes, Cooper and Rhodes 1978), Free Disposal Hull (Deprins, Simar and Tulkens 1984), Linearized Free Disposal Hull (Hall and Simar 2002), Polynomial Estimation (Hall, Park and Stern 1998), Local Linear Fitting (Hall and Park 2004), Moment Type Estimation (Daouia et al. 2010), Pickands Type Estimation (Dekkers, Einmahl and De Haan 1989; Daouia et al. 2010), Conditional Tail Index Estimation (Daouia et al. 2010), Nonparametric Frontier Regularization (Daouia, Florens and Simar 2012), Local Constant Estimation (Gijbels and Peng 2000), Local Extreme-Value Estimation (Gijbels and Peng 2000), and Nonparametric Kernel Boundary Regression (Noh 2014).

[8]They are less sensitive because they estimate a partial frontier well inside the cloud of data points, but near its (empirical) boundary, in such a way to be sensitive to the magnitude of the extreme valuable observations, while also resistant to their influence in case they are entirely error. For technical particulars, see Daouia and Gijbels (2011) and Daouia, Florens and Simar (2012).

Method	AOC	2.5%	50%	97.5%
QR	0.474	0.458	0.49	0.531
Kernel	0.445	0.392	0.449	0.504
SFA	0.448	0.366	0.43	0.526

Table 2: AOC estimates for the simulated data example.

not attempt to model it. Instead, NFMs estimate the density of the observations most relevant to the boundary, ignore the inefficiencies of individual observations in the estimation process, and only compute relative efficiencies *after* the boundary is drawn. In so doing, they combine the strengths of data envelopment with an estimation technique that allows for stochasticity around the boundary.

In the interest of space, we describe the most flexible option, kernel estimation (Noh 2014; Daouia, Noh and Park 2016).[9] Kernel estimation produces a smooth frontier that envelops all the data and minimizes the area under the frontier curve (Noh 2014). Unlike DEA and FDH, this smoothed frontier resists distortion by extreme observations because each observation is weighted by its estimated effect on the underlying boundary. However, because it is a smoothing technique, kernel estimation requires the selection of an appropriate bandwidth. Bandwidth selection is informative because bandwidth and curve complexity are inversely related. Therefore, the analyst can use the selection process to learn about the complexity of the underlying frontier, as well as let the data inform the complexity of the estimated curve. Noh (2014, 507–510) implements a bandwidth selection technique based on the Bayesian Information Criterion, shows that the criterion works well, and demonstrates the superiority of kernel estimation in simulation studies.

But Can't You Always *Draw Boundaries?* A skeptic may wonder whether the boundaries estimated above are trivial.[10] All data have upper and lower boundaries, so there is no obvious distinction between symmetric and asymmetric hypotheses in that regard. There is, however, a meaningful distinction to

[9] For a review of additional DEA-based techniques in the context of analyzing necessary conditions see Dul (2016). While Dul argues that his "Necessary Conditions Analysis" framework (NCA) and fsQCA should be used together to study necessary conditions, he misses the broader applicability of regression-based techniques to the study of asymmetric hypotheses more generally.

[10] This concern is similar to Braumoeller and Goertz (2000) who distinguish trivially necessary conditions from non-trivial ones.

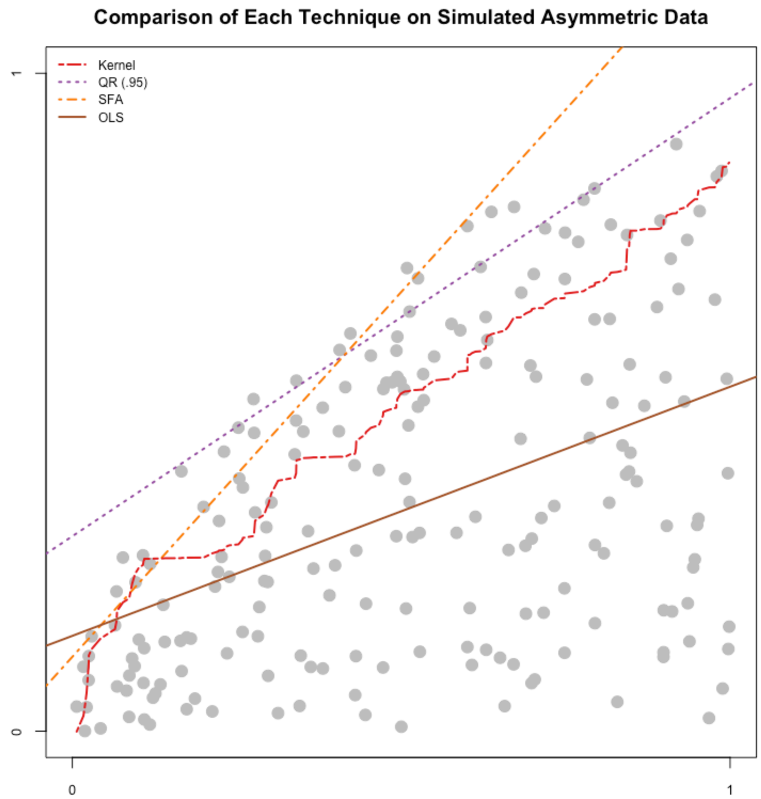

Figure 2: We estimate SFA, QR, and kernel estimation on simulated data. We include OLS to demonstrate that estimating the central tendency of data does not estimate its boundary.

be had based on the extent to which a hypothesized boundary precludes observations: data consistent with an asymmetric hypothesis should be rare or nonexistent in either the upper-left or the lower-right quadrant of a standard scatterplot. In other words, an upper or lower boundary will always exist, but that boundary is trivial if it does not preclude observations across a significant range of the data space.

Accordingly, we introduce a new criterion—Area Outside of the Curve (AOC)—to distinguish boundaries that are meaningful from those that are trivial. AOC is the ratio of the area outside of the estimated boundary (above for lower-triangular relationships, below for upper-triangular ones) to the total area of the scatterplot, defined as $(max(x) - min(x)) \times (max(y) - min(y))$. We calculate AOC for each technique using the simulated data example in Figure 2, as well as report the 95% bootstrapped confidence intervals (5000 replicates). For the simulated asymmetric data, nearly half of the area is devoid of observations, regardless of the technique used. In practice, we expect AOC to be much smaller. This test is not meant to be overly restrictive. Rather, it is meant to weed out relationships that are not especially asymmetric. There is no obvious cutoff value between asymmetric and symmetric relationships, so based on our own simulations we propose what we hope is a fair test: to be considered a nontrivially asymmetric relationship, the AOC should be greater than or equal to 0.2 with 95% certainty, based on the estimated confidence intervals.

In this section we have provided three different options for estimating data boundaries in asymmetric hypotheses. The techniques share two common tasks. First, the analyst must estimate with some degree of uncertainty the location of a single, fixed frontier. Second, the distance from a given observation (X_i, Y_i) to the frontier (in any direction) is unknown, and thus must be estimated. SFA and QR draw stochastic boundaries yet require specifying a particular functional form. The added benefit of accounting for noise around the boundary comes with the cost of either restricting the distribution of the errors and inputs (SFA) or specifying the quantile through which the boundary is drawn (QR). Nonparametric frontier models provide an alternative approach, ranging from the most canonical (DEA and FDH) to more flexible techniques (kernel smoothing). Within the data envelopment framework, this flexibility comes at the cost of potential sensitivity to extreme data as well as leaving inefficiency unmodeled.[11]

Figure 2 shows how each technique performs on simulated asymmetric data. We include OLS as a point of comparison to reinforce the point that regression with an interaction term does not model the boundary of the data. While each technique is capable of estimating data boundaries, they are not equally suitable in all scenarios. Kernel estimation is well suited to applications with small samples because it is less sensitive to outliers. SFA works best when one has more data, is willing to subscribe to distributional assumptions, and is particularly interested in estimating the inefficiency of individual observations. Finally, QR is especially useful for comparisons across space or time because it is less sensitive than other techniques to substantial outliers that can make comparisons difficult. While the estimation is not limited by sample size, one must trade off the requirement to specify the quantile *ex*

[11] Recent innovations in production analysis suggest that Bayesian machine learning techniques using support vector machines are a promising way forward. Scholars hypothesize that these techniques will improve frontier estimation in the presence of measurement error through an adaptive learning process. For further details, see Poitier (2010) and Poitier and Cho (2011).

ante. In the next section, we apply each technique to an example for which it is suited.

EXAMPLES

We now turn to applications of our three techniques for evaluating asymmetric hypotheses—quantile regression, stochastic frontier analysis, and kernel estimation—to three examples of asymmetric hypotheses in the existing literature.[12] We show how asymmetric hypotheses can be investigated in three different scenarios. In addition to estimating the boundary, we include 95% confidence intervals around the boundary based on 5,000 bootstrap replicates. These bounds allow us to show how uncertainty varies with technique and sample size.

The Domestic Sources of Transnational Security Cooperation

Koenig-Archibugi (2004)'s investigation of the relationship between domestic political preferences and transnational security cooperation in the European Union offers a good example of an asymmetric hypothesis that would typically receive attention from qualitative scholars due to its small sample size. It also allows us to demonstrate the utility of kernel estimation in estimating the frontier of upper-triangular relationships. We chose kernel estimation for this example because it works well in small samples, despite being nonparametric, and as such it is the least sensitive to extreme observations.

In addressing the puzzle of why states would surrender political sovereignty by coordinating their foreign and security policy with other EU member states, Koenig-Archibugi argues that "attitudes toward supranational integration are shaped by distinct conceptions of sovereignty and political authority that prevail in the political culture of the member-states" (Koenig-Archibugi 2004, 167). Accordingly, he finds that these attitudes are shaped by a combination of material variables (such as political and economic power and regional governance institutions), as well as ideational variables (including measure of popular and elite-level European identity and policy conformity with other European states).

[12] Replication materials for these examples can be found at the *Political Analysis* dataverse. See, Rosenberg, Knuppe and Braumoeller (2017).

To investigate the domestic political sources of support for supranationalism, Koenig-Archibugi employs fsQCA and discovers the following asymmetric relationship: the combination of strong regional governance (based on the degree of constitutional federalism, the presence of special territorial autonomy, the role of regions in central government, and the existence of regional elections) and policy conformity (having preferences on specific policy issues that are consistent with the preferences of most other EU member states) are sufficient to generate national support for the European Union's Common Foreign and Security Policy (CFSP).

In discussing the substantive implications of this relationship, the author argues that

> a pluricentric constitutional culture removes an important conceptual obstacle—the idea of national sovereignty as unitary and indivisible—from the transfer of decisional powers to the European level. The anticipation that most decisions made in supranational fora would correspond to the government's substantive policy preferences adds to this permissive factor a positive incentive to support supranationalization, and it is the combination of permissive and positive conditions that makes this particular conjuncture sufficient for the outcome. (Koenig-Archibugi 2004, 164)

The author limits his investigation to the original thirteen members states of the European Union. This limited universe of cases makes fsQCA an ideal candidate for examining the possible combinations of conditions that produce supranational security cooperation.

Figure 3 is a replication of Koenig-Archibugi's findings, including a boundary line derived from kernel estimation and bootstrapped confidence intervals. The x-axis reflects fuzzy-set membership scores for EU member states along the intersection of regional governance and policy conformity (two of the six conditions coded in Figure 2 of the original article). The y-axis measures percent support for the EU's Common Foreign and Security Policy (CFSP). The upper-triangular distribution of the data reflects a sufficiency plot, because each membership in the fuzzy set of supranationalist governments (y-axis) is at least as high as a state's membership in the intersection of regional governance and policy conformity (x-axis) (Koenig-Archibugi 2004, 162-163).

We evaluate this asymmetric hypothesis using kernel estimation for three reasons. First, the OLS residuals from SFA models often violate the distributional assumptions of the inefficiency term, and this leads to unreliable inferences even when models are properly specified (Almanidis and Sickles

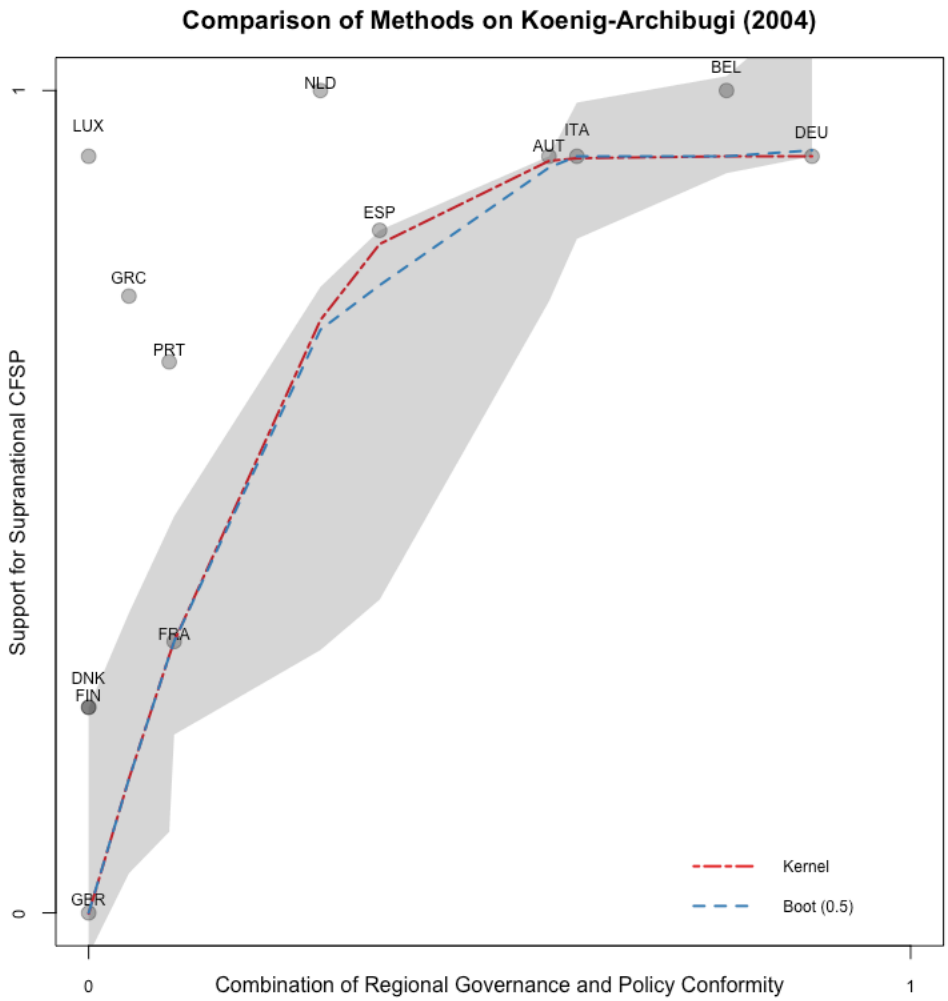

Figure 3: Replication of Koenig-Archibugi (2004) using kernel estimation. The uncertainty around the boundary was generated with 5000 bootstrap replicates.

2011, 204). Second, we do not use QR because the .95 quantile regression would be drawn through only two observations. This coarseness would throw away informative features of the data that a nonparametric technique can expose. Finally, we choose kernel estimation over other nonparametric techniques because it is the least sensitive to extreme observations.

After proceeding, we conduct an AOC test to verify the non-trivialness of this boundary. This exercise may not be necessary because the original analysis was done using fsQCA. Therefore, all observations must fall above the $Y = X$ line by construction. However, we conduct an AOC test as construct validity for our measure. As one can see in Table 3, the area outside the kernel boundary is

method	AOC	2.5%	50%	97.5%
Kernel	0.725	0.475	0.719	0.809

Table 3: AOC estimate for Koenig-Archibugi (2004), with 95% bootstrapped confidence intervals (1000 replicates).

quite large, indicating a strong asymmetric relationship. Despite the small sample size, the area outside the curve does not drop below 0.475 in the bootstrap replications.

We report the kernel estimation boundary in Figure 3. As one can see, the kernel boundary implies a diminishing effect of regional governance and policy conformity on support for supranational foreign and security policy. Even though the kernel estimation boundary is curved, its substantive interpretation remains unchanged from the original analysis—both the United Kingdom (lower left) and Germany (upper right) exhibit ideal-typical relationships between regional governance and policy conformity, and support for transnational security cooperation. Yet, our approach adds considerable value by showing the nonlinear relationship between regional governance, policy conformity, and the lower boundary on support for a supranationalist CFSP. This result can drive future work on what makes Spain, Austria, and Italy important cases for driving transnational security cooperation in the EU.

However, the bootstrapped confidence intervals show the potential dangers of using a statistical technique on an application with so few observations. Because we resample from 13 observations (5000 replicates), the variance in the boundary estimates is quite large. However, the median bootstrap, reported in blue, corresponds well to the original boundary. Despite wide variance at the margin, this similarity indicates that the boundary estimates are rather stable, even when resampling from such a small universe of cases. In general, this demonstration shows that kernel estimation provides a flexible technique for estimating asymmetric boundaries in studies with small samples.

Exploring the Health–Wealth Nexus

Our second example of an asymmetric hypothesis comes from the voluminous literature on the relationship between economic status and health outcomes. This example shows the efficacy of QR for

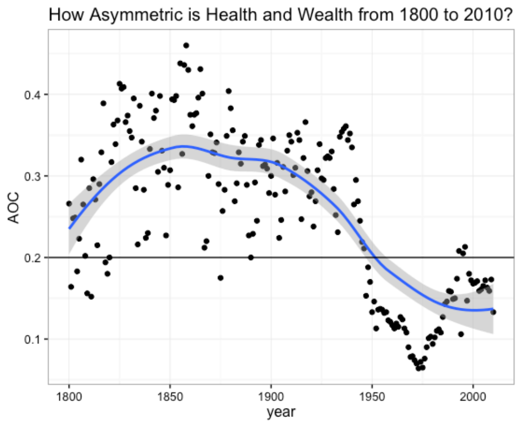

Figure 4: How AOC for the health-wealth example changes over time.

Year	AOC	2.5%	50%	97.5%
1810	0.285	0.209	0.271	0.331
1830	0.355	0.268	0.346	0.393
1850	0.307	0.256	0.309	0.334
1870	0.351	0.275	0.356	0.416
1890	0.200	0.146	0.199	0.252
1910	0.342	0.240	0.330	0.396
1930	0.330	0.292	0.337	0.400
1950	0.133	0.071	0.136	0.181
1970	0.079	0.049	0.081	0.117
1990	0.149	0.124	0.154	0.189
2010	0.133	0.104	0.133	0.158

Table 4: AOC estimate for the Health-Wealth analysis over time, with 95% bootstrapped confidence intervals (1000 replicates).

	Before 1950	After 1950
Prop. AOC > 0.2	0.76	0.02
Mean AOC	0.31	0.13

Table 5: The proportion of years in which the AOC is above the 0.2 threshold. Prior to 1950, the health-wealth relationship was almost always asymmetric. Since 1950 the relationship is much less asymmetric.

evaluating asymmetric hypotheses with many observations over time. In comparison, nonparametric techniques are less suitable for boundary analysis over time because because sensitivity to outliers makes boundaries difficult to compare. We first demonstrate when wealth and health outcomes are asymmetrically related, prior to showing how this relationship changes over time. We conclude by reflecting on the substantive reasons why these patterns emerge.

Following Bergh and Nilsson (2010), we test the claim that GDP per capita (purchasing power parity-adjusted) provides an upper limit on life expectancy. Our data cover the period from 1800-2015, so we are able to extend previous analyses by seeing the degree to which the upper boundary evolves.[13] Once we establish that this is a proper asymmetric relationship, we will be able to see how the region of low-GDP/high-life expectancy changes over time.

[13] Our data come from the aggregated sources on Gapminder. For more details, please see the accompanying replication materials.

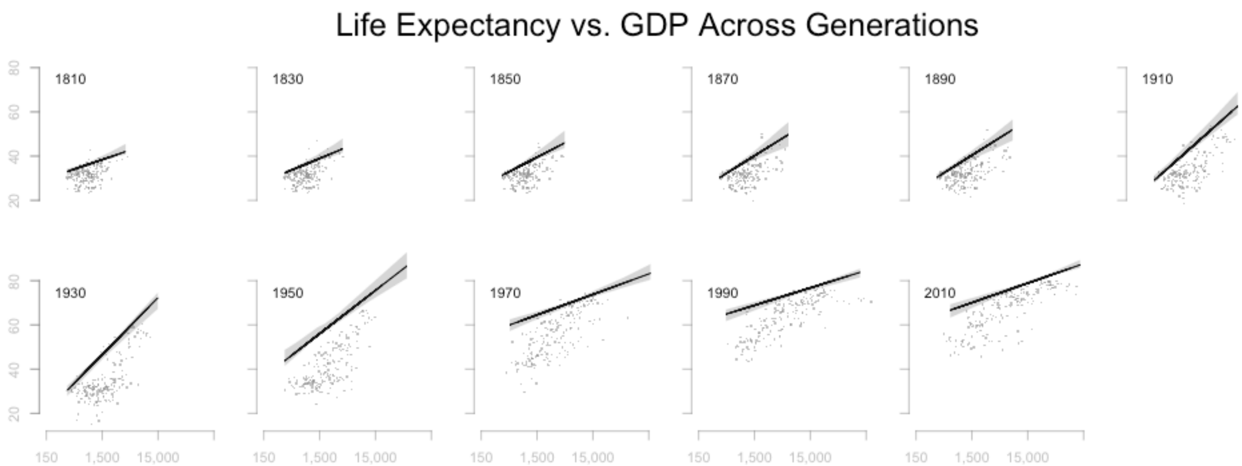

Figure 5: Wealth and the Limits of Human Longevity

We re-estimate the QR model through the .95 quantile for each twentieth year from 1800 to 2015 and plot the resulting frontiers in Figure 5. We then use AOC to see if this relationship is asymmetric. This task is less straightforward when one analyzes time series data because an asymmetric hypothesis in one period may become symmetric in another. In Table 5, we report the AOC for each time period, as well as its 95% bootstrap confidence intervals. In periods prior to 1950, the relationship between health and wealth can be characterized as an asymmetric hypothesis because the AOC is typically above 0.2. However, after 1950, the relationship becomes more symmetric. This evolution becomes starker in Table 5, where one can see the dramatic differences between years prior to and after 1950.

The QR boundaries are presented in Figure 5. We can see that, predictably, the upper limit that GDP places on life expectancy has risen dramatically over the last two centuries. In particular, the increase in life expectancy from 1920 to 1980 was spectacular, mostly due to advances in medicine and public health.

When viewed over time, the slopes of the boundaries are also substantively interesting. When the slope is flat, the upper boundary on life expectancy is more or less the same at all levels of GDP, but a steep line indicates that a state's level of GDP has a pronounced limiting effect on the life expectancy of its citizens. Early in the 19th century, the slopes of the boundaries were much flatter than they became from mid-century on, and only recently have they gotten less steep. Roughly speaking, innovations in health care began to have a noticeable effect as early as 1890 in the richest countries, but it was not until about 1950 that the benefits of those innovations began to improve life expectancy in the poorest countries. In 1930, the upper limit of life expectancy in rich countries like the United States and Germany was about 60 years. By contrast, the upper limit of life expectancy in poorer countries like China and India was about 30 years. By the mid-1990s, however, China had achieved life expectancy nearly on a par with the United States (around 73 years).

In some cases, liberated colonies inherited elements of colonial healthcare systems. In our data, Uganda and Tanzania fit into the category of former colonial states where life expectancy outpaces development. In others, however, healthcare innovations diffuse and are adapted by large, developing states. This is most evident in China, where life expectancy surged between 1950 and 1970 (despite the murderous Great Leap Forward in the early 1960s). While both sets of cases provide examples of life expectancy outpacing economic development, the explanation behind this relationship operates via different causal mechanisms.

In addition, these factors help explain why the relationship between GDP and life expectancy became less asymmetric over time. Prior to the innovations and development of the mid-20th century, it is clear that in the absence of economic development, life expectancy was low. In other words, rather than predict life expectancy, GDP placed an upper limit on it. This relationship is the definition of an asymmetric hypothesis. However, since the 1950s, it is clear from looking at AOC and Figure 5 that the relationship has become more symmetric. This change suggests that the barriers to longevity in the poorest countries had been successfully pushed back, though substantial variation among poorer countries remains.

Party Size, Electoral Systems, and Democracy

Finally, we reexamine the findings on party system size that were presented by Clark, Gilligan and Golder (2006). CGG explore the hypothesis that the number of parties in a given country (party system size) is a function of the interaction of electoral system permissiveness and social heterogeneity. Theoretically, these two dimensions limit the maximum number of parties, producing a frontier that is of great importance to scholars of comparative electoral systems. In Duverger's (1963) canonical presentation, the author argues that single-member district plurality (SMDP) electoral systems are sufficient to produce a two-party system—or equivalently, that a non-SMDP electoral system is necessary to produce a multiparty system. Neto and Cox (1997) give these hypotheses more specific form by observing that "[t]he effective number of parties in a polity should be a multiplicative rather than an additive function of the permissiveness of the electoral system and the heterogeneity of the society." CGG examine a simplified version of Neto and Cox's model in order to demonstrate that much of the heterogeneity in party system size can be explained by social (in this case, ethnic) heterogeneity.[14] However, the authors' interactive model fails to capture the intuition that, when the absence of data is the phenomenon of interest, our goal should be to delineate the boundary at which this absence begins. Interaction terms are incapable of doing this because OLS estimates the central tendency of the data.

In this example, we show how SFA can be used to estimate the upper boundary of the data. Our stochastic frontier model retains the interactive essence of the CGG model, while estimating the stochastic upper bound of the data rather than their central tendency. SFA also allows us to model heterogeneity by making the magnitude of the inefficiency term a function of the same variables that define the frontier. The results in Figures 6 and 7 demonstrate how party system size can be understood as a boundary phenomenon rather than a matter of central tendency. The stochastic party system size frontier is described by a surface of the form:

$$\begin{aligned} Frontier = 2.997 - 0.330 \times \text{Electoral System Permissiveness} - 1.056 \times \text{Social Heterogeneity} \\ + 1.347 \times \text{Electoral System Permissiveness} \times \text{Social Heterogeneity} \end{aligned}$$

This frontier describes the maximum size of party systems we should expect for given levels of electoral system permissiveness and social heterogeneity. In addition, the AOC for upper percentile of the frontier is 0.206, suggesting that this hypothesis is asymmetric (see Table 6).

One ought to note that, unlike other regression models, in SFA "parameter estimates by themselves

[14] Further details can be found in Clark, Gilligan and Golder (2006, 323).

Method	AOC	2.5%	50%	97.5%
SFA	0.330	0.206	0.597	0.872

Table 6: AOC estimate for CGG (2006), with 95% bootstrapped confidence intervals (1000 replicates).

have little informational value; they are simply means to the end of producing a frontier" (Fenn et al. 2008, 94). This feature of SFA reinforces our point that studying asymmetric hypotheses does not reduce to null hypothesis significance testing. T he frontier gives specific estimates for how the upper boundary of party system size is associated with various levels of electoral system permissiveness and social heterogeneity. The frontier also indicates how individual cases perform relative to their theoretical upper limit.

For this reason, scholars most often use SFA results to describe the "inefficiency" of individual cases—that is, how far the value of the dependent variable in a given case falls below the theoretical maximum dictated by the values of the independent variables. Substantively, these results show that some polities (Figure 7) are notably more "inefficient" than others at producing the expected number of parties given their electoral system and degree of social heterogeneity. Consider the case of Bolivia, the least "efficient" observation in the sample. Factors such as indigenous politics, labor dynamics within the tin and copper mining community, or political divisions between voters in the Andes versus populations in the Amazon basin may promote *sui generis* party dynamics which are otherwise non-ideal from a theoretical point of view (Whitehead 1981). Similar inefficiencies may exist in Colombia, where coffee, gold, and coca cultivation affect regional party formation, mostly through effects on local conflict dynamics (Richani 2013).

Our re-analysis of CGG follows the lead of the set theorists in finding the boundary between the region in which observations occur and the region in which they largely do not (Figure 6). SFA also allows us to harness the statistical rigor of large-N models. In so doing data boundaries tell us something fundamentally different about the social world than regression models that estimate conditional averages. We can now usefully discuss the upper limit of party system size given electoral system size and social heterogeneity, as well as the performance of individual countries. Our results are indicative of how SFA can be used to explore data boundaries as well as how it can improve and respect the

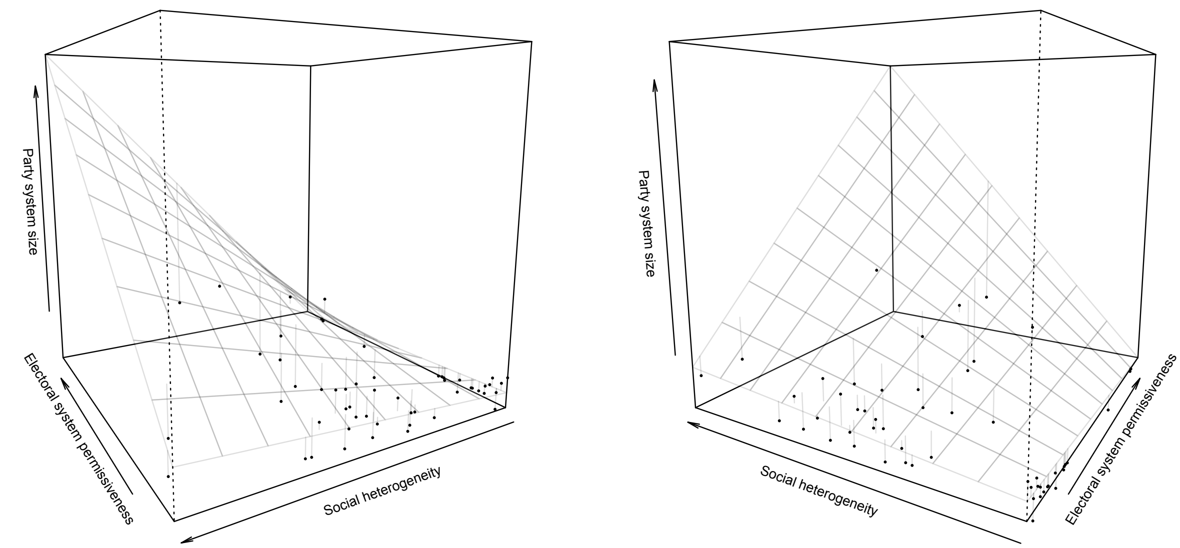

Figure 6: A reanalysis of Clark, Gilligan, and Golder (2006) using stochastic frontier analysis, showing the upper limit of party system size given social heterogeneity and electoral system permissiveness.

approaches of both set-theoretic and statistical methodologists.

CONCLUSION

This article presents both a conceptual clarification of asymmetric hypotheses and a discussion of the methodologies available to explore them. Most existing methods used by political scientists fail to capture the core insight motivating asymmetric hypotheses: the boundaries separating zones of data from areas that lack observations are the primary features of interest. While multiplicative interactive terms are incapable of modeling these boundaries, set-theoretic methods end up assuming a fixed boundary by construction. QR, SFA and NFMs, by contrast, allow us to estimate the boundary as well as to explore the distance between the boundary and specific observations of interests. These techniques neither ensure nor preclude causal relationships. Instead, they equip the analyst to explore asymmetric hypotheses—a feature of the data that may previously have been overlooked.

We also hope to have shown that qualitative and quantitative approaches to asymmetric hypotheses can usefully inform one another and provided a set of methodological tools that will be useful to both.

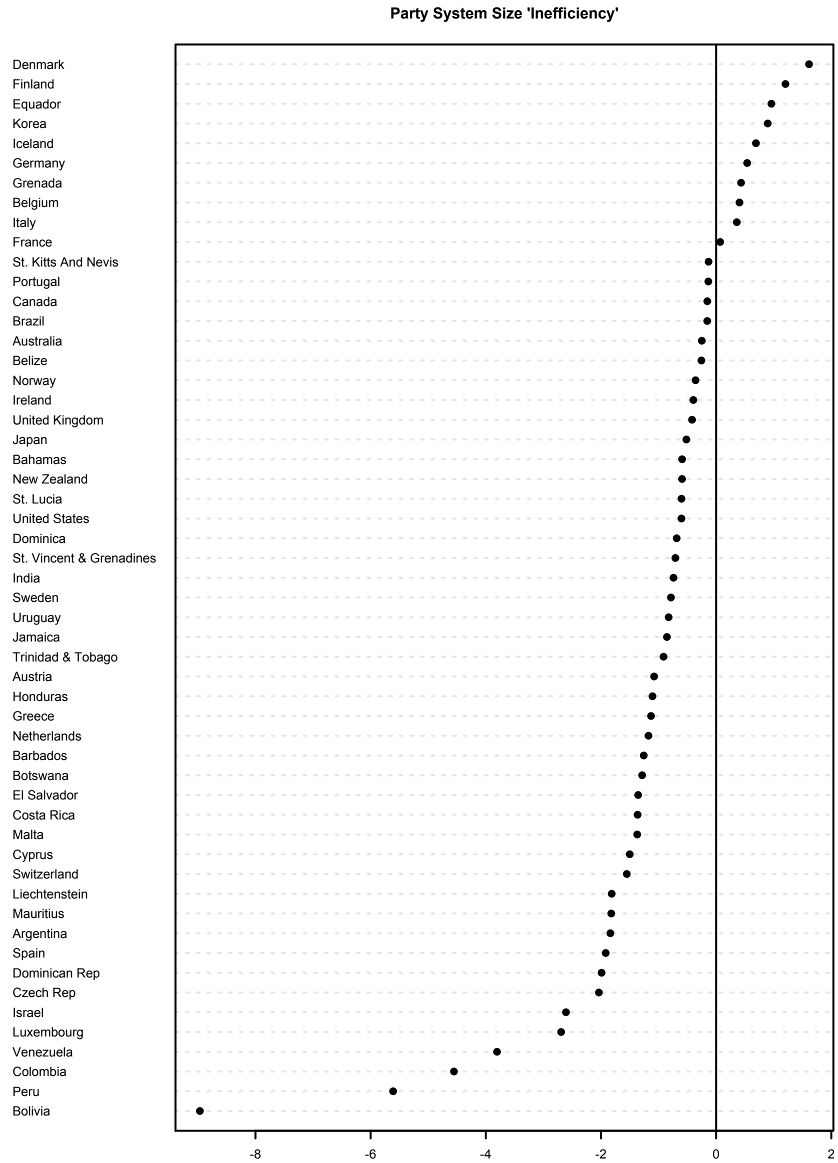

Figure 7: Inefficiencies in Party System Size

In so doing, we hope to have narrowed the schism between qualitative and quantitative scholars and opened the door for more question-driven rather than method-driven research.[15] Such an effort will facilitate continued progress away from the dogmatic “Age of Regression” and toward a future in which scholars test their theories with methods that conform to theoretical logic rather than continuing

[15] We recognize that often differences in methods chosen by qualitative and quantitative scholars are driven by the types of data to which they have access. Therefore, this claim will not apply to all qualitative scholars. The qualitative scholars that will benefit most from these methods are those who assign numeric values to cases and use set-theoretic methods to evaluate asymmetric hypotheses. We thank an anonymous reviewer for emphasizing this point.

to shoehorn a complex social world into a linear-additive straitjacket.

REFERENCES

Abbott, Andrew. 1988. "Transcending General Linear Reality." *Sociological Theory* 6(2):169–186.

Aigner, Dennis, C. A. Knox Lovell and Peter Schmidt. 1977. "Formulation and Estimation of Stochastic Frontier Production Function Models." *Journal of Econometrics* 6(1):21–37.

Almanidis, Pavlos and Robin C. Sickles. 2011. The Skewness Issue in Stochastic Frontiers Models: Fact or Fiction? In *Exploring Research Frontiers in Contemporary Statistics and Econometrics*. Springer pp. 201–227.

Angrist, Joshua D and Jörn-Steffen Pischke. 2008. *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton, NJ: Princeton University Press.

Bergh, Andreas and Therese Nilsson. 2010. "Good for Living? On the relationship between globalization and life expectancy." *World Development* 38(9):1191–1203.

Braumoeller, Bear F. 2015. "Guarding Against False Positives in Qualitative Comparative Analysis." *Political Analysis* 23(4):471–487.

Braumoeller, Bear F. and Gary Goertz. 2000. "The Methodology of Necessary Conditions." *American Journal of Political Science* 44(4):844–858.

Bueno de Mesquita, Bruce. 1983. *The War Trap*. New Haven, CT: Yale University Press.

Bueno de Mesquita, Bruce, James D Morrow, Randolph M Siverson and Alastair Smith. 1999. "An Institutional Explanation of the Democratic Peace." *American Political Science Review* 93(4):791–807.

Charnes, Abraham, William W. Cooper and Edwardo Rhodes. 1978. "Measuring the Efficiency of Decision making Units." *European Journal of Operational Research* 2(6):429–444.

Clark, William Roberts, Michael J. Gilligan and Matt Golder. 2006. "A Simple Multivariate Test for Asymmetric Hypotheses." *Political Analysis* 14(3):311–331.

Daouia, Abdelaati, Hohsuk Noh and Byeong U Park. 2016. "Data Envelope Fitting with Constrained Polynomial Splines." *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 78(1):3–30.

Daouia, Abdelaati and Irène Gijbels. 2011. "Robustness and Inference in Nonparametric Partial Frontier Modeling." *Journal of Econometrics* 161(2):147–165.

Daouia, Abdelaati, Jean-Pierre Florens and Léopold Simar. 2012. "Regularization of Nonparametric Frontier Estimators." *Journal of Econometrics* 168(2):285–299.

Daouia, Abdelaati, Jean-Pierre Florens, Léopold Simar et al. 2010. "Frontier Estimation and Extreme Value Theory." *Bernoulli* 16(4):1039–1063.

Debreu, Gerard. 1951. "The Coefficient of Resource Utilization." *Econometrica: Journal of the Econometric Society* pp. 273–292.

Dekkers, Arnold LM, John HJ Einmahl and Laurens De Haan. 1989. "A Moment estimator for the Index of an Extreme-Value Distribution." *The Annals of Statistics* pp. 1833–1855.

Deprins, D, L Simar and H Tulkens. 1984. Measuring Labor Inefficiency in Post Offices. In *The Performance of Public Enterprises: Concepts and Measurements*, ed. M Marchand, Pestieau P and Tulkens H. North-Holland pp. 243–267.

Dion, Douglas. 1998. "Evidence and Inference in the Comparative Case Study." *Comparative Politics* 30(2):127–145.

Dul, Jan. 2016. "Necessary Condition Analysis (NCA) Logic and Methodology of "Necessary but Not Sufficient" Causality." *Organizational Research Methods* 19(1):10–52.

Duverger, Maurice. 1963. *Political Parties: Their Organization and Activity in the Modern State*. New York: Wiley.

Farrell, Michael James. 1957. "The Measurement of Productive Efficiency." *Journal of the Royal Statistical Society. Series A (General)* 120(3):253–290.

Fenn, Paul, Dev Vencappa, Stephen Diacon, Paul Klumpes and Chris O'Brien. 2008. "Market Structure and the Efficiency of European Insurance Companies: A Stochastic Frontier Analysis." *Journal of Banking & Finance* 32(1):86–100.

Gartzke, Erik. 1998. "Kant We All Just Get Along? Opportunity, willingness, and the origins of the democratic peace." *American Journal of Political Science* 42:1–27.

George, Alexander L. and Andrew Bennett. 2005. *Case Studies and Theory Development in the Social Sciences*. Boston, MA: MIT Press.

Gijbels, Irène and L Peng. 2000. "Estimation of a Support Curve via order Statistics." *Extremes* 3(3):251–277.

Goertz, Gary, Tony Hak and Jan Dul. 2013. "Ceilings and Floors Where Are There No Observations?" *Sociological Methods & Research* 42(1):3–40.

Greene, William H. 2008. The Econometric Approach to Efficiency Analysis. In *The Measurement of Productive Efficiency and Productivity Growth*, ed. Harold O. Fried, C. A. Knox Lovell and Shelton S. Schmidt. New York: Oxford University Press pp. 92–250.

Hall, Peter and Byeong U Park. 2004. "Bandwidth Choice for Local Polynomial Estimation of Smooth Boundaries." *Journal of Multivariate Analysis* 91(2):240–261.

Hall, Peter, Byeong U Park and Steven E Stern. 1998. "On Polynomial Estimators of Frontiers and Boundaries." *Journal of Multivariate Analysis* 66(1):71–98.

Hall, Peter and Léopold Simar. 2002. "Estimating a Changepoint, Boundary, or Frontier in the Presence of Observation Error." *Journal of the American Statistical Association* 97(458):523–534.

Hug, Simon. 2013. "Qualitative Comparative Analysis: How Inductive Use and Measurement Error Lead to Problematic Inference." *Political Analysis* 21(2):252–265.

Kaufman, Stuart J., Richard Little and William C. Wohlforth. 2007. *The Balance of Power in World History*. New York, NY: Palgrave Macmillan.

Koenig-Archibugi, Mathias. 2004. "Explaining Government Preferences for Institutional Change in EU Foreign and Security Policy." *International Organization* 58(01):137–174.

Koenker, Roger. 2005. *Quantile Regression*. New York, NY: Cambridge University Press.

Koopmans, Tjalling C. 1951. "Efficient Allocation of Resources." *Econometrica: Journal of the Econometric Society* pp. 455–465.

Krasner, Stephen D. 1983. *International Regimes*. Ithaca, NY: Cornell University Press.

Krogslund, C., D. D. Choi and M. Poertner. 2015. "Fuzzy Sets on Shaky Ground: Parameter Sensitivity and Confirmation Bias in fsQCA." *Political Analysis* 23(1):21–41.

Li, Yuhui and Matthew S. Shugart. 2016. "The Seat Product Model of the Effective Number of Parties: A Case for Applied Political Science." *Electoral Studies* 41:23–34.

Lieberson, Stanley. 1985. *Making it Count: The Improvement of Social Research and Theory*. Berkeley, CA: University of California Press.

Lijphart, Arend. 1990. "The Political Consequences of Electoral Laws, 1945–85." *American Political Science Review* 84(02):481–496.

Lupia, Arthur and Mathew D. McCubbins. 1998. *The Democratic Dilemma: Can Citizens Learn What They Need to Know?* New York: Cambridge University Press.

Maoz, Zeev and Bruce Russett. 1993. "Normative and Structural Causes of Democratic Peace, 1946–1986." *American Political Science Review* 87(03):624–638.

Meeusen, Wim and Julien Van Den Broeck. 1977. "Efficiency Estimation From Cobb-Douglas Production Functions With Composed Error." *International Economic Review* 18(2):435–444.

Mercer, Jonathan. 1996. *Reputation and International Politics*. Ithaca, NY: Cornell University Press.

Moore, Barrington. 1966. *Social Origins of Dictatorship and Democracy: Lord and Peasant in the Making of the Modern World*. Boston: Beacon Press.

Morgan, Stephen L. and Christopher Winship. 2014. *Counterfactuals and Causal Inference*. 2 ed. New York: Cambridge University Press.

Neto, Octavio Amorim and Gary W. Cox. 1997. "Electoral Institutions, Cleavage Structures, and the Number of Parties." *American Journal of Political Science* 41(2):149–174.

Noh, Hohsuk. 2014. "Frontier Estimation using Kernel Smoothing Estimators with Data Transformation." *Journal of the Korean Statistical Society* 43(4):503–512.

Poitier, Kerry R. 2010. *True Efficient Frontiers in Production Performance Assessment Using Machine Learning and Data Envelopment Analysis*. Miami, FL: University of Miami.

Poitier, Kerry and Sohyung Cho. 2011. "Estimation of True Efficient Frontier of Organisational Performance Using Data Envelopment Analysis and Support Vector Machine Learning." *International Journal of Information and Decision Sciences* 3(2):148–172.

Ragin, Charles C. 1987. *The Comparative Method: Moving Beyond Qualitative and Quantitative Strategies*. Berkeley, CA: University of California Press.

Ragin, Charles C. 2000. *Fuzzy-set Social Science*. Chicago, IL: University of Chicago Press.

Ragin, Charles C. 2008. *Redesigning Social Inquiry: Fuzzy Sets and Beyond*. Chicago, IL: University of Chicago Press.

Ray, Subhash C., Subal C. Kumbhakar and Pami Dua. 2015. *Benchmarking for Performance Evaluation: A Production Frontier Approach*. New York, NY: Springer.

Richani, Nazih. 2013. *Systems of Violence: the Political Economy of War and Peace in Colombia*. New York: SUNY Press.

Rosenberg, Andrew S., Austin J. Knuppe and Bear F. Braumoeller. 2017. *Replication data for: Unifying the Study of Asymmetric Hypotheses*. Harvard Dataverse.
URL: *http://dx.doi.org/0.7910/DVN/9NXGHP*

Russett, Bruce. 1994. *Grasping the Democratic Peace: Principles for a post-Cold War World*. Princeton, NJ: Princeton University Press.

Russett, Bruce, Christopher Layne, David E. Spiro and Michael W. Doyle. 1995. "Correspondance: The Democratic Peace." *International Security* 19(04):164–184.

Schimmelfennig, Frank. 2001. "The Community Trap: Liberal Norms, Rhetorical Action, and the Eastern Enlargement of the European Union." *International Organization* 55(01):47–80.

Seawright, Jason. 2005. "Assumptions, Causal Inference, and the Goals of QCA." *Studies in Comparative International Development* 40(1):39–42.

Seawright, Jason. 2016. *Multi-Method Social Science: Combining Qualitative and Quantitative Tools*. New York, NY: Cambridge University Press.

Seiford, Lawrence M. 1996. "Data Envelopment Analysis: the Evolution of the State of the Art (1978–1995)." *Journal of Productivity Analysis* 7(2-3):99–137.

Simar, Léopold. 2007. "How to Improve the Performances of DEA/FDH Estimators in the Presence of Noise?" *Journal of Productivity Analysis* 28(3):183–201.

Simar, Léopold and Paul W. Wilson. 2013. *Estimation and Inference in Nonparametric Frontier Models: Recent Developments and Perspectives*. Boston, MA: Now.

Simar, Léopold and Valentin Zelenyuk. 2011. "Stochastic FDH/DEA Estimators for Frontier Analysis." *Journal of Productivity Analysis* 36(1):1–20.

Taagepera, Rein and Matthew Soberg Shugart. 1993. "Predicting the Number of Parties: A Quantitative Model of Duverger's Mechanical Effect." *American Political Science Review* 87(02):455–464.

Talmadge, Caitlin. 2015. *The Dictator's Army: Battlefield Effectiveness in Authoritarian Regimes*. Ithaca, NY: Cornell University Press.

Thiem, Alrik, Michael Baumgartner and Damien Bol. 2016. "Still Lost in Translation! A Correction of Three Misunderstandings Between Configurational Comparativists and Regressional Analysts." *Comparative Political Studies* 59:742–774.

Tsebelis, George. 1999. "Veto Players and Law Production in Parliamentary Democracies: An Empirical Analysis." *American Political Science Review* 93(03):591–608.

Whitehead, Laurence. 1981. "Miners as Voters: the Electoral Process in Bolivia's Mining Camps." *Journal of Latin American Studies* 13(2):313–346.

Wimmer, Andreas. 2014. "Nation Building. A Long-Term Perspective and Global Analysis." *European Sociological Review* 31(1):30–47.

Yu, Keming and Rana A Moyeed. 2001. "Bayesian Quantile Regression." *Statistics & Probability Letters* 54(4):437–447.