Advancing the Empirical Research on Lobbying

John M. de Figueiredo* and Brian Kelleher Richter**

Prepared for the Annual Review of Political Science 2014

November 2013

ABSTRACT

This essay identifies the empirical facts about lobbying which are generally agreed upon in the literature. It then discusses challenges to empirical research in lobbying and provides examples of empirical methods that can be employed to overcome these challenges—with an emphasis on statistical measurement, identification, and casual inference. The essay then discusses the advantages, disadvantages, and effective use of the main types of data available for research in lobbying. It closes by discussing a number of open questions for researchers in the field and avenues for future work to advance the empirical research in lobbying.

Keywords: Lobbying, Interest Groups, Empirical Methods

* jdefig@duke.edu, Duke University, School of Law and Fuqua School of Business, and NBER, 210 Science Drive, Durham, NC 27708-0360.
** brian.richter@mccombs.utexas.edu, University of Texas at Austin, McCombs School of Business, 2110 Speedway, B6500, CBA 5.250, Austin, TX 78712.

We thank David Austen-Smith, Jordi Blanes i Vidal, Charles Cameron, David Primo, David Lewis, Brian Roberts, Tim Werner, and an anonymous referee for helpful comments on earlier drafts of this paper.
I. INTRODUCTION

One of the central tenets of representative democracy is the right of individuals, by themselves or in groups, to petition elected officials and the government. These petitions are designed to influence the opinions, policies, and votes of legislators and other government officials. One outgrowth of this right to petition the government has been the creation and evolution of organized interest groups comprised of individuals, companies and other organizations. These organized interests employ a variety of methods to influence government policies including campaign contributions, endorsements, grassroots campaigns, media campaigns, and lobbying. A robust literature has spawned hundreds of papers on each of these topics in political science.

While empirical papers written on the subject of campaign contributions and money have dominated the statistical work on the influence of interest groups in politics (Ansolabehere et al. 2003), there has emerged over the past decade a growing literature in political science and related disciplines that empirically examines lobbying. This renewed interest seems to find its roots in three areas. First, newly created disclosure rules on lobbying and lobbyists’ effort coupled with more innovative data collection methods have led to the creation of a number of new datasets on lobbying that are now available to researchers, mostly, but not exclusively, focused on activity in the United States. Second, the rise of these datasets has created a keen public interest in the process and statistical regularities in lobbying efforts. Third, political scientists have very recently joined forces with economists to create a more seamless research boundary on the topic between the disciplines, jointly developing more advanced and better identified statistical models of lobbying.
In this paper, lobbying is defined as the transfer of information in private meetings and venues between interest groups and politicians, their staffs, and agents. Information takes the theoretical representation of a message, and in practice, may have many forms: statistics, facts, arguments, messages, forecasts, threats, commitments, signals or some combination of the aforementioned. Interest groups have budgets for and spend money on these activities, but that money is not transferred explicitly to politicians (as it is with campaign contributions) (de Figueiredo 2002). If we assume, following most of political economy literature, that a politician’s objective function is comprised of re-election to the current office, promotion to higher office, and ideological pursuits, then the politician seeks information on how her position on a given issue or issue set will affect these outcome variables (Milyo 2001). There may be intermediate forms of information—such as how many jobs a policy position will create, how will my constituents be affected by a particular yea or nay vote, whether business leaders will support me in the next election, etc. —but ultimately, the key piece of information the politician cares about is understanding how position-taking on various issues will affect her re-election, promotion, and ideological policy outcomes.¹

This essay synthesizes four main aspects of the empirical lobbying literature. First, we summarize the generally accepted findings in recent empirical and statistical advances in the informational lobbying literature. Second, we discuss statistical methods that we believe are

¹ This essay does not discuss a number of areas. The essay does not cover the campaign finance literature or illegal lobbying expenditures—such as situations where lobbying becomes money transfers. In those situations, lobbying expenditures are bribes and most appropriately discussed with respect to the literature on corruption. This paper does not examine how interest groups employ direct-to-voter media campaigns and other forms of public pressure campaigns (e.g. Hall & Reynolds 2012). The paper does not cover the literature on the structure and value of political connections (e.g. Faccio 2006; Goldman et al. 2009) outside of the purely lobbying context. Within the lobbying literature, the essay does investigate legislative lobbying but does not explicitly pursue the large literature on lobbying of bureaucratic agencies that deserves its own analysis. Papers in this latter area include Naoi & Krauss (2009); de Figueiredo & Tiller (2001); Holburn & Vanden Bergh (2004); McKay (2011); de Figueiredo & Kim (2004); Yackee & Yackee (2006), McKay & Yackee (2007), and McKay (2011) to name only a few.
particularly fruitful in obtaining statistical identification and making causal inference in the lobbying literature. Third, we discuss the new datasets that have recently become available in the field and what particular advantages and disadvantages each type of dataset engenders in a research program. Finally, we provide what we believe are some possible avenues for future research. Ultimately, this essay examines what we know about lobbying, what we’d like to know about lobbying, and how we might make headway in finding the answers.

II. EMPIRICAL REGULARITIES IN LOBBYING

This section takes the large number of empirical studies on lobbying and attempts to distill a few central facts where there seems to be consensus in the literature. We attempt to stay away from equivocal findings and focus, instead, on findings where there is broad agreement.

A. WHO LOBBIES, HOW MUCH THEY LOBBY, AND HOW LOBBYING IS ORGANIZED

The first regularity in the data is that lobbying is pervasive in the American political system and seems to be quite important in the political systems of many other developed countries. Milyo et al (2000) shows that lobbying expenditures at the U.S. federal level are five times that of political action committee (PAC) campaign contributions. Moreover, the relative magnitude of lobbying expenditures to interest groups’ campaign contributions continues to persist at similar levels. In 2012, the amount spent by organized interests on lobbying the federal government amounted to $3.5 billion annually—while the estimated amount spent on campaign contributions by interest groups’ political action committees, super-PACs, and 527 organizations was approximately $1.55 billion over the two-year 2011-to-2012 election cycle, or approximately $750 million annually (Federal Election Commission 2013; Center for
Responsive Politics 2013). Thus, even in the past few years, lobbying expenditures remained approximately five times interest group campaign finance contributions.

There is now overwhelming evidence to support a second general regularity in the data: corporations and trade associations comprise the vast majority of the lobbying expenditures by interest groups. Lobbying expenditures by corporations and trade associations represent over 84% of total interest group lobbying expenditures at the U.S. federal and 86% of total lobbying expenditures at the state level (de Figueiredo 2004). In contrast, this same paper finds issue-ideology membership groups represent 2% of lobbying expenditures at the federal level and 7% of lobbying expenditures at the state level. In addition, large organized interest groups and groups that are supported by large corporations are more likely to lobby than smaller groups and groups that are supported by smaller corporate interests. This is true for a wide of variety of firms across a number of industries and a number of years in the United States (Ansolabehere et al. 2002; Hansen et al. 2004; Hochberg 2009; Richter et al. 2009; Guo 2009; Hill et al. 2011). Moreover, in particular industries and issue areas, such as tariffs and trade disputes, this same regularity holds (Bombardini & Trebbi 2006; Schuler 1996; Lee & Baik 2010). Non-profits, such as universities, also exhibit this same pattern—larger universities are more likely to lobby that smaller ones (de Figueiredo & Silverman 2006). Finally, these patterns are not confined to North America, but have been shown to be true in other developed and transition economies (Naoi & Krauss 2009; Sukiassyan & Nugent 2011).

While businesses represent a substantially large proportion of total lobbying expenditures, they represent a smaller proportion (but a still a majority) of the number of interest groups lobbying. In a survey of ninety-eight randomly selected issues before the Congress in 1999-2001, Baumgartner et al (2009) found that trade and business associations, business
corporations, professional associations, and coalitions specific to the issue represented 54% of all lobbying groups. Citizen groups, unions, foundations, think tanks, governments, institutions, and other groups represented the remainder. At the state level, Lowery and Gray (1996) found that approximately 30% of lobbying groups registered in the American states were governments or social groups. This data combined with the expenditure data above suggests that business groups’ lobbying expenditures are, on average, higher than non-business interests.

A third regularity in the data is that large corporations and well-funded groups are more likely to lobby independently than smaller groups. Most small interest groups are more likely to lobby using only trade associations. Richter et al. (2009) and Kerr et al. (2013) found that only 10% of publicly-traded firms actually lobbied on their own behalf. Moreover, these authors show that groups that lobby tend to show substantial persistence and serial correlation in their lobbying efforts—with Kerr et al. (2013) reporting that the probability that a firm lobbies in a given year conditional on lobbying in the prior year is 92%. Various authors have suggested the reason for this is that smaller interests lack the resources to front the fixed costs for a lobbying organization, they lack the necessary access to politicians to influence public policy, they do not individually carry the political power to influence outcomes, or they have issues that arise only intermittently (Bertrand et al. 2012). Companies, however, may also avoid using external lobbyists when corporate secrets and innovations are at risk (de Figueiredo & Tiller 2001), when political systems differ (Mahoney 2007), and when free-riding or issue characteristics lend themselves to collective as opposed to individual efforts (Bombardini & Trebbi 2012). The use of trade associations by smaller organizations and companies may mask these smaller organizations’ participation in lobbying, as trade associations in the U.S. must report lobbying efforts, but trade association members do not generally have to report the fees they pay to
become members of those organizations, and thus these fees are not tabulated as lobbying expenditures by smaller firms. Harstard & Svensson (2012) have gone further to show that small companies may actually be more likely to engage in bribes as a substitute to lobbying. Together, this implies that in a snapshot of lobbying in an industry or issue area, it should be common to see both individual interest groups and trade associations lobbying at the same time and larger lobbying groups having offices present near the center of political decision-making (Hansen & Mitchell 2000; Campos & Giovannoni 2007; Schuler et al. 2002).

A fourth regularity in the data is that lobbying increases when the issues are more relevant or salient, or there are big stakes for the organized interest (Leech et al. 2005; Bonardi 2005; Baumgartner et al. 2011, Baumgartner & Leech 2002; Calderia et al. 2000). A related finding is that budgeting drives increases in lobbying efforts. Leech et al. (2005) show at the federal level that issues and agencies with larger budgets draw more lobbying effort by interest groups. de Figueiredo (2014) and de Figueiredo & Cameron (2012) exploit the cross-sectional and time-series variation in budgeting rules and budget size in 38 U.S. states over many years and political regimes to show that budgeting time periods result in a 19% increase in lobbying expenditures by interest groups. Just as issue salience draws interest group attention, so do budget issues, government monies, and tax issues (La Pira et al. 2012).

B. THE NUMBER OF INTEREST GROUPS REGISTERED

In 1996, Gray & Lowery authored The Population Ecology of Interest Representation on the organizational ecology of interest groups in American states. Since that time there has been an explosion in the number of papers that examine the ecology of interest groups at the state level, focusing on both their density and diversity based primarily on counts of groups rather than dollars spent. The argument put forward in the book was first articulated in biology and
then extended into the social science literature by the sociologists Hannan & Freeman (1978, 1984) and with advancements by Carroll & Hannan (1992). The core argument put forward in the Hannan-Freeman-Carroll theory is that organizations in a population are buffeted by two forces: legitimation and competition. Gray & Lowery (1996) brought this framework to interest groups and collected data on the names and counts of all the registered lobbying groups in each state in certain specified years (1975, 1980, 1990, and subsequently 1997). Putting the legitimation and competitive effects together results in an inverted U-shaped curve of the number of organizations over time. Initially, the legitimation effect causes a rise in the number of organizations in a field. Then the competitive effect takes over, causing a shakeout of the number of organizations. The number of organizations finally stabilizes around a steady-state number. More recent advances in this area consistently show a second regularity, that the number of interest groups is correlated with the size of economy and number of issues the legislature faces (Baumgartner et al. 2009; Gray & Lowery 1998) and positively correlated with number of constituents and constituent interests (Berkman 2001; Lowery & Gray 1998). Some of the most recent findings in the field are particularly focused on health care lobbying (Benz et al. 2011; Lowery et al. 2005).

C. EXPERTISE, CONNECTIONS, TARGETING, AND COUNTERACTING

One of the key questions researchers have attempted to tackle, with limited success, is whether lobbyists derive value from what they know (expertise) or who they know (connections). Empirically we know that at least some lobbyists tend to specialize in issues. Lobbying groups like Cassidy Associates or Sidley Austin specialize in lobbying on behalf of academic institutions and telecommunications issues, respectively, suggesting issue expertise is valuable. Measuring the value of expertise is difficult. Cameron & de Figueiredo (2013)
develop a model based on expertise and test its implications using state level panel lobbying data. They find substantial empirical support in the intensity and targeting of lobbying for a model based on expertise, but do not explicitly measure expertise nor estimate its value.

Bertrand et al. (2012) also find indirect support for the expertise explanation, demonstrating that lobbyists who specialize in particular issues are more likely to access politicians of an opposite party, presumably because the politicians value the issue expertise of the lobbyist.

However, a recent paper by Blanes i Vidal et al. (2012), discussed in-depth later in this essay, examines the lobbying revenue of former Senatorial staffers who have become lobbyists. The paper investigates what happens to the revenues when there is a change in focal Senator’s status. They find a 23% decline in a lobbyist’s revenue after the Senator on whose staff the lobbyist formerly served is defeated in an election or retires from Congress. Likewise, Bertrand et al. (2012) also find that lobbyists tend to follow the politician to whom they are connected, even when that politician switches committees and handles substantially different issue areas. This would suggest a “who you know” story drives at least a good proportion of lobbying revenues. As we discuss at the end of this essay, the “who you know” versus “what you know” debate is an attractive area for empirical research in lobbying.

Despite this ongoing debate, there is consensus on a number of aspects of targeting in the lobbying literature. First, powerful legislators are most likely to be targeted for lobbying. These legislators usually have agenda setting power as sponsors or co-sponsors of bills (Hojnacki & Kimball 1999), are on issue-relevant or generally powerful and influential committees, like Appropriations, Budget, or Finance Committees (Drope & Hansen 1998; Honjacki & Kimball 2004; Duso 2005), or in the congressional leadership, such as committee chairs and ranking members or majority or minorities leaders (Evans 1996).
The second set of regularities in the studies on targeting focus on where targets sit on the ideological or position continuum. There seems to be a growing consensus in the literature that both allied (Heberling 2005; Honjacki & Kimball 1999; Kollman 1997; Calderia & Wright 1998; Hall & Deardoff 2006, Hall & Miler 2008) and marginal legislators on both sides of the issue (Bertrand et al. 2012; Gawande et al. 2012; Keheller & Yackee 2009; Holyoke 2003; Tung 2011) are targeted for lobbying efforts, but staunch enemies are not targeted by interest groups. Because the current U.S. lobbying disclosure regulations do not require lobbyists to identify which legislators they are targeting, the papers on the subject usually rely on survey data or inference from lobbying expenditure data. A clear linking of theory to the empirical regularities on targeting is found in the counteractive lobbying literature by Austen-Smith & Wright (1994, 1996). In empirical work done on Supreme Court nominations, Austen-Smith & Wright find empirical support for the theory that lobbyists target marginal legislators to “swing” them to the lobbyist’s position, and they target friends to “counteract” the lobbying from opposition groups. Hall & Miler (2008) and Hojnacki & Kimball (1998) have critiqued the counteractive lobbying approach, instead arguing that the interest groups’ legislative allies are the primary targets, followed by marginal legislators. They show empirically in accordance with the theoretical predictions of Rotemberg (2003) and Hall & Deardoff (2006), that lobbyists will target mainly allies (as a legislative subsidy) and agenda-setters to influence the shape of legislation, or to encourage these allies to in turn lobby marginal and influential policymakers. These two approaches to targeting paint a picture of lobbying where a variety of different legislators are targeted, even by one group, based on their position in Congress and their position on the issue.
D. THE EFFECTIVENESS OF LOBBYING

A final area of interest to scholars is understanding and quantifying how effective interest group lobbying is in obtaining policy or other outcomes. This is an extraordinarily challenging question to tackle because econometric identification is problematic and causal mechanisms are extremely difficult to isolate. Because of this, the significant number of papers measuring the effectiveness of lobbying must overcome a number of challenges that make statistical inference, estimation, and interpretation difficult. Literature in this area should be viewed with a critical eye toward these drawbacks by the reader. (This point is discussed further in the next two sections of this essay.)

The most common area for researchers to examine is international trade. A variety of papers have attempted to show that lobbying affects tariffs, customs classifications, and dumping determinations across a wide variety of countries (Grasse & Heidbreder 2011; Gawande et al. 2006, 2012; Stoyanov 2009; Ehrlich 2008; Drope & Hansen 2004; Baylis & Furtan 2003; Tung 2011; Tavares 2006; Lee & Baik 2010). A second area that has received substantial attention over the past five years is the effect of lobbying on a variety of financial and regulatory accounting issues. These papers purport to demonstrate that lobbying affects the ability of firms to influence financial regulation and legislation, to engage in revenue hiding and avoid fraud detection, and to garner excess returns in the marketplace through a variety of mechanisms (Mian & Trebbi 2011; Blau et al. 2013; Hochberg et al. 2009; Igan et al. 2012; Yu & Yu 2011; Gelbach 2006; Richter 2011; Hill et al. 2011; Adelino and Dinc 2013). A third area that has received attention substantial attention is the effect of lobbying on appropriations and budgeting. A set of papers in this vein argue that lobbying is effective in allowing interest groups and their allies to obtain a larger slice of government budget and contracts in a number of different
countries (Evans 1996; Helland 2008; Kelleher & Yackee 2009; Alt et al. 1999; de Figueiredo & Silverman 2006, 2007, details discussed in further depth in the next section). A fourth area that has received increasing attention is the effect on lobbying on taxation. Richter et al. (2009) has found that firms that lobby are more likely to pay lower income taxes to the U.S. federal government. (This paper is also discussed in further depth in this essay.) Schone et al. (2013) finds that tax development credits are more likely to be granted in the presence of lobbying in France. Finally, there have been papers measuring the effect of lobbying on a variety of different outcomes. These papers attempt to show that lobbying affects judicial confirmations (Caldeira & Wright 1998; Austen-Smith & Wright 1996), immigration policy (Facchini et al. 2011), trade association entry barriers (Morris & Neeley 2001), regulated prices in telecommunications (Duso 2005), technology diffusion (Comin & Hobijn 2009), the general passage of bills (Grasse & Heidbreder 2011), and overall economic performance (Horgos & Zimmerman 2009), to mention just a few areas.

Perhaps the most comprehensive academic study on the effects of lobbying was carried out by Baumgartner et al (2009). In an eight-year study covering ninety-eight policy issues before Congress, the authors track in detail lobbying efforts by nearly 2,200 advocates; they also follow the results of lobbying efforts and policy outcomes up to four years after the policies were proposed. They find that both sides of an issue are usually able to mobilize relatively equal amounts of resources. One reason for this is any side of an issue is usually comprised of a heterogeneous mix of corporate, citizen, and government advocacy groups that bring to bear resources comparable to the opposition. Baumgartner et al (2009) find that policies are usually very stable and resilient to change, but when changes do occur, the policy changes tend to be substantial.
Overall, there have been numerous studies attempting to estimate the effects of lobbying on policy outcomes. However, the validity of their results depends crucially on the dataset and econometric methods employed to cleanly identify and isolate the causal effect of lobbying. In the following sections we discuss the data and these methods for the empirical analysis of lobbying.

E. SUMMARY

The burgeoning empirical literature on lobbying has created many new findings. There is now an emerging consensus around a set of facts. Lobbying is a pervasive institution in the American political landscape with lobbying expenditures representing five times the dollar volume of interest group campaign contributions. This activity is most likely to be pursued by large firms and interest groups. Large firms will have a higher probability of lobbying independently than small firms, while small firms will tend to agglomerate their lobbying in trade associations. Budgeting, highly salient issues, and issues which impact groups more will attract more lobbying effort. In addition, the number of interest groups is positively correlated with macroeconomic activity, constituent interest, and legislative workflow. These facts are well established and there are few returns to future researchers demonstrating these facts are true, yet again.

There is also a set of facts that has been shown to be true, but the underlying reasons for the regularities in the data are a bit more opaque. The politicians targeted by interest groups tend to be powerful allied agenda setters and marginal legislators on the issue, but tend not to be staunch opponents. While this has been demonstrated, there are a number of different theories as to why this might occur. Recent literature has found strong evidence that “who you know” matters and there is some evidence that issue expertise may also matter to targeting. However,
the precise magnitude and mechanism underlying this fact is not well understood. Finally, there are a large number of papers that show lobbying has returns. However, these papers employ a variety of different empirical methods, methods that each have drawbacks to isolating and identifying the causal effects of lobbying. In the next section, we discuss these methodologies in further detail.

III. EMPIRICAL APPROACHES TO STUDYING LOBBYING

Advances in social science research methods since the 1990s have allowed researchers who studying lobbying to move away from correlational studies by applying quasi-experimental research designs that allow stronger causal inferences to be made. These advances now permit scholars to avoid, as Baumgartner & Leech (1996) note, the “pitfalls of one-shot cross-sectional designs,” but do not obviate the need for understanding the underlying narratives and institutional details of lobbying. In fact, understanding how lobbying works in practice is arguably even more important than before in implementing the design of empirical research.

In this section, we outline how advances in research design and statistical methods can enable scholars to provide more reliable answers to otherwise difficult questions. We outline the challenges inherent in empirical research on lobbying, outline research designs and statistical approaches to overcome these challenges, and provide examples of recent research that has employed these approaches. While each of the articles we will highlight has slightly different attributes and research designs, they all adhere to a common set of practices that we believe are necessary to advance the next generation of empirical research in lobbying.
A. CHALLENGES TO CONDUCTING EMPIRICAL ANALYSIS IN LOBBYING

In order to advance our understanding of these and other areas, however, researchers are encountering a number of statistical challenges that are making causal inference difficult. The first challenge that has arisen is the time series persistence and stickiness of lobbying effort and lobbying registrations within interest groups. With little within-interest-group variation in lobbying over time, it can be difficult to draw inferences from a panel dataset without external shocks that affect different groups at different times (as opposed to all groups at once). For example, imagine a certain interest group allocates $100,000 to lobbying per year and is awarded a government contract of $1M every year. While this relationship may be suggestive, we cannot describe the causal relationship between lobbying and contracts without a shock to the group’s budget or to the contract’s size as, holding everything else constant, the variation necessary for techniques like fixed effects is not present.

The second challenge of empirical studies in lobbying is a rather significant omitted variable problem. For instance, consider that multiple instruments for exercising political influence may be used in tandem but only some of them may be observed. If this is the case, we may falsely attribute an outcome to an interest’s lobbying efforts when the effect is really due to the interest’s location in a key district. Given the nature of lobbying and some actors’ desire to operate under the radar, researchers are faced with omitted variables even when using the best available data. Data sets may not contain data on observable factors we would like to include as controls in the analysis—or there may be factors that are simply unobservable such as interests’ or lobbyists’ innate ability at lobbying. The omitted variable problem can be extremely problematic if these omitted variables are correlated with the error term in a regression, as this will result in biased parameter estimates and incorrect causal inference. We believe this is a
pervasive problem in the lobbying literature. As quantification of outcomes, rather than
descriptions of phenomena, becomes an increasingly important goal, avoiding omitted variable
bias will become a primary concern.

A third challenge for the literature is endogenous selection into the lobbying process.
The decision to lobby by an interest group is not a random event—and hence does not meet the
idealized world of an experimental trial, in which some interest groups would be assigned to
lobby and others not. Not permitting random selection could lead to biased results because the
group assignment process could be correlated with outcomes. An interest group’s decision to
lobby is likely driven by the group’s expected reactions of other groups who may also lobby and
by the focal group’s anticipated outcomes. That is, groups are more likely to lobby when they
believe they are more likely to succeed. Endogenous selection will result in incorrect statistical
inference and biased parameter estimates in a standard regression model. Moreover, endogenous
selection can make the direction of causality difficult to assess, since interests may only lobby
the way they do because they expect a benefit from doing so. We note this problem could be
pervasive given that far fewer interests choose to lobby in practice than could (about 10%). If
the majority of interests that do not lobby choose not to because they expect no return, then
without properly accounting for selection, estimates of returns will be biased upwards.
Moreover, the exact level of lobbying activity, the venues targeted, etc. may be chosen in such a
way as to optimize on the outcome of the interest, further complicating causal stories about
lobbying linked to outcomes. This is related to the omitted variable problem, and like it, makes
newer questions that require causal inference difficult.

Finally, recent empirical work has attempted to make the links between theory and data
more tightly coupled. As the empirical lobbying literature began to emerge, it was largely
independent of theory. However, with tighter linking of theory and empirical work expected going forward, researchers face the challenge that though theories about lobbying are often about information, the information (or message) is, in most datasets, generally not directly observable. Hence, empirical work must develop mechanisms that were not required in the past to test theories of informational lobbying when the information cannot be observed or might be incomplete.

B. EMPIRICAL METHODS FOR ADVANCING EMPIRICAL LOBBYING RESEARCH

In this section, we highlight some of the research designs and empirical methods that help to overcome the challenges identified, providing an example of each approach. The approaches covered here are difference-in-differences, event studies, instrumental variables, selection models, and structural modeling.

1. Difference-in-Differences Approach with Exogenous Shock

The first research design that allows researchers to identify causal effects of lobbying activities is the difference-in-differences approach. Differences-in-differences is a technique whereby one measures the change in the treatment and control groups across pre- and post-treatment periods. The main advantage of the differences-in-differences approach vis-à-vis the challenges presented above is that it exploits an exogenous source of variation to deal with the stickiness or persistence issue in lobbying data. A secondary advantage of the approach is that, through the use of fixed effects in panel data, it eliminates concerns about all time-invariant observation-unit specific omitted variables.
Blanes i Vidal et al. (2012) employ this difference-in-difference with exogenous shocks approach in an examination of revolving door lobbyists. Blanes i Vidal et al. (2012) quantify how valuable ‘who you know’ is to the lobbyist. Focusing on former congressional staffers who are contract lobbyists for interest groups, they examine how the revenues of lobbyists change when politicians to whom the lobbyists are connected via prior employment retire or are defeated in an election. Because many politicians exit over time, the authors have a key source of exogenous shock or variation that allows the authors use to identify how valuable “who you know” is.

Empirically, they estimate the equation:

\[ R_{it} = \beta P_{it} + X'_{it} \cdot \theta + \alpha_i + \gamma_{t}^{pc} + \epsilon_{it} \]

where \( R_{it} \) represents the revenue the lobbyist earns in a specific period, \( P_{it} \) is an indicator of whether or not the lobbyist’s former employer is an active politician or not, \( X_{it} \) represents time-varying observable lobbyist attributes, \( \alpha_i \) controls for unobservable individual lobbyists specific factors that remain time-invariant, and \( \gamma_{t}^{pc} \) controls for unobservable time-specific factors depending upon the party and chamber to which a lobbyist is connected.

Because a key source of variation that the research design exploits for identification is prior employment with a particular politician, the authors can rely on politician exits to generate their statistical results. By having a relatively long time-dimension panel (22 periods) and including in their specifications time-varying lobbyist attributes, lobbyist fixed effects, and party-chamber specific time fixed-effects, Blanes i Vidal et al. are able to attribute drops in lobbyist revenue vis-à-vis the within-lobbyist counterfactual trend to lobbyists losing someone who they know. Hence, with the differences-in-differences approach, the authors can identify
that who lobbyists know matters, accounting for approximately 23% of the value of a given lobbyist’s services.

2. Event Study Approach

The second method that can be utilized to overcome some of the challenges identified is the event study approach. This quasi-experimental approach relies on an exogenous shock to one group to allow comparison between a control and experimental groups in the pre- and post-shock conditions. The exogenous events that shock the lobbying systems help avoid concerns about endogenous selection into a particular behavior. Like the differences-in-difference approach, event studies also are well suited to handling the stickiness or persistence issue in lobbying since they focus on how actors’ behavior changes (or does not change) in response to a major shock to the system.

We highlight Jayachandran (2006) who uses a financial market event study to examine how firms that align themselves with politicians from the majority party may benefit from that position. Although this paper is not directly about lobbying, it nicely executes and illustrates an approach that can be usefully employed in lobbying studies. In May 2001, Senator James Jeffords announced he was leaving the Republican Party to become an Independent, making the Democrats the majority party in the U.S. Senate, in turn, altering the political landscape as Democrats took over key leadership roles and now had more power over the legislative agenda.

Jayachandran (2006) employs a financial market event study around Jeffords’ switch to construct a counterfactual of how the stock prices of certain firms fared versus how they would have fared had Jeffords not left the majority Republican Party. She then examines whether

---

2 Other papers that profitably use event study analysis to analyze political outcomes include Roberts (1990), Werner (2011), and Hillman et al. (1999).
aberrations between the actual and counterfactual performance could be predicted by firms’ relative relationships with both parties. To construct the abnormal returns—or the difference between actual and counterfactual market performance of firms—around Jeffords’ announcement, Jayachandran employs two steps. First, she calculates expected returns for firms by estimating the following equation using data from before the event-period:

$$Rtn_{it} = \alpha_i + \beta_i MktRtn_{it} + \epsilon_{it}$$

Then, she assumes that the estimates of $\hat{\alpha}_i$ and $\hat{\beta}_i$ will persist around the event, allowing her to calculate abnormal returns (as the difference between actual and expected returns) during the event window (i.e. the days around Jefford’s announcement) as:

$$AbnRtn_{it}^{\text{event}} = Rtn_{it}^{\text{event}} - (\hat{\alpha}_i + \hat{\beta}_i MktRtn^{\text{event}})$$

Jayachandran then investigates whether deviations in firms’ actual performance and counterfactual performance can be explained by firms’ alignment with the majority party who unexpectedly loses power. To do so, she regresses estimates of firms’ political party alignment prior to Jefford’s announcement (as proxied by soft money contributions) on her estimated abnormal returns. More formally, she estimates:

$$AbnRtn_{it}^{\text{event}} = \beta_0 Dem_i + \beta_R Rep_i + X_i' \cdot \theta + \epsilon_{it}$$

After controlling for other observable firm-specific factors that might explain why the market performance of some firms deviated from their long-run trend, Jayacahndran finds stronger alignment with the former majority party, i.e. the Republicans, hurt the firm’s market performance. Hence, the paper implements an event-study based research design that allows the author to uncover the importance of structuring political relationships in a way that aligns
interest groups with politicians who have agenda setting abilities and other forms of legislative power, while remaining flexible to adapt should the political environment change. While Jayachandran focuses directly on campaign contributions rather than informational lobbying, the approach in her work illustrates how event-studies are potentially fruitful quasi-experimental approaches for lobbying.

3. Instrumental Variables Approach

A third method for identifying the substantive impact of lobbying is the implementation of an instrumental variables approach. This approach relies on identifying a variable that is correlated with the endogenous explanatory variables, conditional on other covariates, but is uncorrelated with the error term in the regression equation. The instrumental variables approach is primarily focused on solving the challenge of omitted variable bias and endogeneity of right hand side variables.

de Figueiredo & Silverman (2006) employ this method in their study of the effects of lobbying on earmarks granted to academic institutions by Congress from 1997 to 1999. The goal in the paper is estimate the returns to university lobbying by measuring the size of the earmarks Congress appropriates to an institution as a function of that institution’s lobbying expenditures. One reason this setting is chosen for empirical analysis is that other common forms of political action that might also contribute to political outcomes, such as grassroots organization of employees for political purposes and political action committees, is not permitted by non-profit charitable organizations such as universities, allowing the authors to isolate the effect of lobbying. Empiricists might be tempted to estimate the following equation:

\[ \text{Earmarks}_i = \beta \text{Lobbying}_i + X'_i \cdot \theta + \alpha + \epsilon_i \]
Unfortunately, this approach would be invalid if the level of lobbying that an institution chooses is a function of its expected return on lobbying efforts in the form of earmarks. The authors overcome this very common challenge in estimating the effects or outcome of lobbying effort by finding an instrumental variable to use in a two-stage least squares research design (Angrist et al. 1996).

de Figueiredo & Silverman argue that overhead rates charged to universities as part of grant funding are a valid instrument because overhead rates (i) are a meaningful cost shifter for universities in using grant monies, (ii) are the result of negotiations between each university and the bureaucratic government agencies that disperse funds, and (iii) are not under the purview of elected politicians who are lobbied to insert earmarks into legislation. Hence, higher overhead rates should cause universities to invest more in lobbying, but should not directly result in earmarks. Rather, the earmarks appear through the higher lobbying investments that universities make because of higher overhead rates.

In the first stage, they estimate the determinants of lobbying as a function of overhead rates among other factors. Specifically they estimate:

\[ Lobbying_i = \beta Overhead_i + W'_i \cdot \theta + \alpha + \epsilon_i \]

where \( Overhead_i \) is their key instrumental variable that affects the level of lobbying a university chooses, but does not affect the size of earmarks it receives, and where \( W_i \) represents other observable factors which could affect how much a university spends on lobbying.

In the second stage, they estimate the causal effect of lobbying on obtaining earmarks by using the instrumented value of lobbying estimated in the first stage regression, \( \hat{Lobbying}_i \), as their key independent variable. Hence, they estimate a version of:
\[ Earmarks_i = \gamma \text{Lobbying}_i + X'_i \cdot \zeta + \alpha + \eta_i \]

where \(X_i\) represents other observable factors which could affect the earmarks a university receives.

Because the authors have an instrument in their empirical setting that determines lobbying levels, but is likely not correlated with the error term in the regression, they are able to estimate the direct effect of lobbying on that outcome free of simultaneity bias—overcoming one of the key challenges to estimates of the efficacy of lobbying on outcomes. de Figueiredo & Silverman find that lobbying has a significant effect on the size of earmarks received when the legislator representing the district of the university is on an appropriations committee in Congress, but that lobbying has no statistically significant independent effect on earmarks when the legislator representing the university is not on an appropriations committee. Thus, the paper also enhances our understanding of when lobbying has a payoff.

4. Selection Model Approach

A fourth method of addressing the challenges of estimating the effects of lobbying is to utilize selection models. In this approach, the non-random assignment of subjects to two groups (e.g. control and experimental) is explicitly modeled. This method is geared toward resolving challenges related to groups choosing to lobby on a non-random basis. When coupled with panel data methods, they can also be used to reduce concerns about time-invariant, unit-specific omitted variables bias.

Richter et al. (2009) employ a selection model in their study of the relationship between lobbying efforts and the effective tax rates companies pay as a check on whether or not there is
an endogenous process of selection into lobbying on taxes. Nevertheless, the authors argue that because all firms prefer lower effective tax rates to higher effective tax rates, endogenous selection into lobbying should not be a problem for that outcome, noting that only 10% of firms actually lobby.

Simple panel methods, without accounting for the selection effect, would estimate the following equation:

\[ ETR_{it} = \beta \text{Lobbying}_{it-1} + X'_{it-1} \cdot \theta + \alpha_i + \alpha_t + \epsilon_{it} \]

where \( ETR_{it} \) is firm’s \( i \) effective tax rate at time \( t \), \( \text{Lobbying}_{it-1} \) is the firm’s lobbying expenditure in the previous time period, \( X_{it-1} \) are lagged time-varying observable factors, \( \alpha_i \) and \( \alpha_t \) are firm and time fixed effects. In the presence of sample selection bias, this equation would suffer from omitted variable bias and the coefficient on lobbying expenditures, \( \beta \) would result in incorrect inference about the effects of lobbying on tax rates.

To check for this problem, Richter et al. implement a two-step Heckman (1979) selection model that accounts for any selection effect. In the first step, a binary selection equation is estimated—and the coefficients from it are used to calculate an estimate of the inverse Mills ratio \( (\lambda) \). This represents an estimated selection hazard for the probability that a given firm selected into lobbying.

In Richter et al.’s (2009) case, they estimate the decision to lobby or selection equation as:

\[ \text{If}_t \text{Lobby}_{it} = F(\omega W_{it-1} + \alpha + \mu_{it}) \]

where \( \text{If}_t \text{Lobby}_{it} \) is a binary variable representing the decision to lobby or not, \( F( . ) \) represents the cumulative density function of a standard normal distribution since they are estimating a
probit model and $W_{it-1}$ represents factors that influence a firm's decision to lobby including both factors included in $X_{it-1}$ in the baseline estimation equation of interest and factors not included. Given functional form assumptions, the inclusion of additional factors is not strictly necessary for estimation; however, the inclusion of additional factors that affect selection but not outcomes greatly increase the robustness of the estimates (Sartori 2003). In the case of Richter et al., they use as additional factor liquid assets, such as cash, which when plentiful may make it easier for firms to lobby, but are unlikely to directly affect the effective tax rate.

Richter et al. then use the estimates $\hat{\omega}$ to construct the estimate of the inverse Mills ratio as:

$$\hat{\lambda} = \frac{\phi(\hat{\omega}W_{it-1})}{\Phi(\hat{\omega}W_{it-1})}$$

where $\phi$ represents the standard normal probability density function and $\Phi$, its cumulative density function. This is then inserted into the outcome equation to control for endogenous selection into lobbying to overcome selection bias issues and make causal inference.

$$ETR_{it} = \beta Lobbying_{it-1} + \eta \hat{\lambda} + X_{it-1}' \cdot \theta + \alpha_{it} + \alpha_{t} + \epsilon_{it}$$

Using the sample selection method outlined here, Richter et al. (2009) find not only that increasing lobbying efforts by 1% over a firm’s baseline level in a given year predicts lower effective tax rates by 1.07 percentage points on the margin in the next period, but also that these results are not driven by firms’ decisions to select into lobbying because they covet lower tax rates. This paper and the use of selection models, however, cannot resolve fully whether or not firms, that are already lobbying, opportunistically increase their efforts at opportune moments—making it difficult to say whether or not money is left on the table by firms that do not lobby.
5. Structural Modeling Approach

A final approach that can be used to overcome the inherent challenges in empirical research in lobbying is to employ a structural modeling approach. Although not a quasi-experimental approach like the four presented above, a structural model does build out the econometric equations directly from a theoretical model so that each parameter in the statistical model has theoretical underpinnings. Moreover, the structural estimation approach allows the researcher to calibrate the theoretical model to policy experiments (doing “what-if” analysis) which cannot be done as rigorously with a reduced form approach. The simultaneity and causal direction of effects are derived in the formal model, and the empirical tests match the theoretical model. This minimizes endogenous selection concerns if the theoretical model is believed to reflect the true institutional details of the situation. In doing this, structural models overcome some, but not all of the traditional issues that arise with statistical inference. Underlying these models, though, are often a number of strong assumptions that are required to generate the result, assumptions which may be so strong as to render these models suspect.

Kang (2012) provides a current example of a structural econometrics approach to lobbying. She builds a formal all-pay contest model of a lobbying process in which groups choose to select into a counteractive lobbying environment. Applying this model to energy issues, she estimates the model’s parameter values in examining energy interest groups’ support and opposition for policies embedded within various bills during the 110th Congress. To obtain results, Kang assumes that it is difficult for interest groups to change their lobbying strategies within a single two-year session of Congress.
Kang finds that the effect of energy interest groups lobbying expenditures on a policy’s equilibrium enactment outcome to be very small, but that these small shifts in policy are actually quite profitable relative to the small investments in lobbying firms make. She finds that lobbying in this setting has a marginal return in policy of 140% return. The value of structural econometric approaches comes from their ability to marry theory with data in such a way that predictions from theoretical models can be estimated broadly.

C. SUMMARY OF EMPIRICAL METHODS AND RESEARCH DESIGN

The five examples above demonstrate that recent research in lobbying has been able to overcome a number of empirical challenges inherent in assessing various aspects of lobbying activity by using more advanced empirical methods. While each paper approaches its respective question with a different methodological approach, they almost all share a quasi-experimental design structure that has advantages over more common empirical designs used in lobbying research.³

Each of the articles we have reviewed provides an application of a research design to a different question of lobbying. They each demonstrate that lobbying can be valuable in some contexts, but also show that the value of lobbying depends on timing, the interest groups’ targets, and other factors that appear to play a role in determining the outcomes of lobbying efforts.

³ Despite the attractiveness of these new statistical approaches, these methods are not fool-proof. Differences in differences can suffer from inconsistent standard errors (Bertrand et al. 2004). Event study analysis may have inappropriate windows or confounding factors in the window that are correlated with the outcomes, yielding incorrect causal inference. Weak instruments may eliminate the benefits of an instrumental variables regression (Stock & Staiger 1997). Lack of a convincing selection mechanisms or small samples may doom selection models (Sartori 2003). Unrealistic assumptions might make structural models unbelievable (Angrist & Pischke 2010). Nevertheless, when correctly employed and carefully used, these methods can make substantial strides in addressing the challenges that empirical lobbying research faces.
Beyond the methods used in the examples above, there are also a number of quasi-experimental approaches in other literatures that might become useful in advancing the empirical analysis of lobbying. A commonly used method in other social science fields to help deal with sample selection and omitted variables bias issues is to compare treated units with untreated units who are otherwise similar using advanced matching methods (Rosenbaum & Rubin 1983). Regression discontinuity designs exploiting knife-edge assignments to treatment and control groups help overcome questions about sample selection and omitted variable bias as well (Imbens & Lemieux 2008; Angrist & Lavy 1999). Synthetic controls have been constructed to empirically identify treatment effects in single case studies (Abadie & Grazebal 2003; Abadie et al. 2010). These approaches have yet to be employed in the empirical analysis of lobbying, but will likely yield fruit if used in the appropriate settings in future research.

Despite the importance of methodological approaches to advance the empirical analysis of lobbying, researchers should continue to adhere to the basic tenets of any good research design. These include: 1) clearly defining the research question, 2) understanding the institutional details, 3) understanding the source of variation and the counterfactual, 4) identifying and explaining the mechanisms that explain the results, and 5) eliminating alternative explanations. These factors coupled with the more advanced methodological techniques should allow researchers to push the envelope on empirical research in lobbying by allowing quantifiable causal inference.

IV. DATA

While it is important to understand what types of empirical approaches are effective and likely to yield strong causal inferences about lobbying, implementing such research designs
cannot be conducted without appropriate and sufficiently high quality data. The amount of data available on lobbying is increasing rapidly as interest in the topic grows both within and outside of academia, as disclosure laws proliferate in new jurisdictions, and as time passes allowing existing datasets stemming from older disclosure laws to grow with time. In addition, researchers are becoming more creative in their own data collection efforts. In this section, we review the basic types of data available and the advantages and disadvantages associated with each. We also consider what future sources of data will advance empirical research in lobbying. We stress, though, that advances in measuring lobbying will be most useful in answering the returns to lobbying only to the extent that outcome variables can be well-measured too and clearly linked to lobbying efforts. Too often, the link between lobbying activity and the outcome in questions can be incomplete or tenuous, particularly when authors fail to provide institutional details about the linkages.

A. TYPES OF DATA

Broadly speaking there are three general classes of data typically collected on lobbying activity: surveys, registries, and transaction records. Disclosure rules and survey questions typically dictate what information is available in a particular dataset. The questions asked of survey participants vary from study to study and disclosure laws vary from jurisdiction to jurisdiction. This leads to datasets of varying quality and degrees of usefulness for researchers who want to empirically identify causal outcomes related to lobbying activity rather than simply produce summary statistics or generate partial/conditional correlations. Each class of data has its own strengths and weaknesses as does each distinct dataset.
1. Surveys

Some of the first modern empirical academic research on lobbying dates back to Milbrath’s (1963) survey work and this tradition has continued with surveys and interviews of lobbyists (Heinz et al. 1993; Baumgartner et al. 2009), interest groups (Schlozman & Tierney 1986; Wright 1990; Kollman 1998; Yadav 2008), and bureaucrats (Furlong 1998; Waterman et al. 1998).

One key advantage of survey data is that the survey questions are flexible and allow researchers to investigate topics that lobbying disclosure laws would not permit. For example, Nelson & Yackee (2012) use survey data to study when organizations choose to lobby as a coalition instead of by themselves. This is particularly useful because disclosure laws typically do not require organizations to declare whether or not their lobbying activity is coordinated with other groups as part of a coalition nor do they typically require associations to disclose their members. Survey data can also be particularly useful when conducting cross- or multi-jurisdictional studies, since the disclosure requirements and institutional rules may vary (subtly or dramatically) from location to location—making transactional data from different locations difficult to compare without understanding how the institutional rules shape lobbying behavior. Campos & Giovannoni (2007) use World Bank Survey data to show that lobbying and corruption are substitutes and that lobbying is more likely to occur than bribery in locations with stronger institutions. Yadav (2008) collects her own comparative survey data in India and China to analyze when interest groups get involved in the lobbying process. Hence, survey data is most

---

4 While governments collected some data on lobbying at that time, the information content was limited and it was difficult to access.
likely to be fruitfully collected when observational data on lobbying is limited due to the nature of disclosure laws or when the questions of interest are cross-jurisdictional in nature.

Despite these advantages, survey data has a number of disadvantages. Survey data in lobbying frequently suffers from significant non-random non-response rates, lack of random samples, ex-post recollection, small numbers of observation units, cross-sectional dimensionality only, and limited ability to verify the validity of answers (Groves et al. 2009). All of these issues combined make it impossible for empiricists to fully control for unobservable individual observation-unit fixed characteristics, one of the important characteristics of good statistical work highlighted earlier. Moreover, the problems with surveys will likely generate multiple sources of statistical bias making causal inference difficult if not impossible.

2. Lobbying Registrations

Disclosure laws have also generated public sources of data that are widely available today. Among this data, registries of lobbyists are by far the most common. Nearly every jurisdiction that attempts to regulate lobbying activity has a registry requirement, although the requirements of who must register and what information registrants must provide vary greatly across jurisdiction. Researchers have heavily relied on this registry information to make broad inferences about interest group participation in lobbying activity (e. g. Gray & Lowery 1996; Wolak et al. 2005; Gray et al. 2004).

Registry data offer a number of advantages when studying lobbying. Because the registry data is available for a full population of interest groups over a long period of time, it can help to answer questions about who is registered to lobby. A second advantage of many lobbying registries is that they also contain information on the not just the groups, but the individual
lobbyists who are registered to lobby. This allows researchers not only to track lobbyists over time, but also to link lobbyists to various interest groups. This may be particularly useful to researchers attempting to identify the use of in-house versus contract lobbyists by groups, and could also lead to creation of lobbying network maps that include the connectedness, proximity, and centrality of lobbyists and interest groups (e.g. La Pira et al. 2012).

Registry data, however, suffers from a number of practical drawbacks. Given the differences in registration rules across jurisdictions, it is difficult to make direct comparisons across them. In some geographies registration is voluntary; in others, registration rules differ; and in yet others, there is substantial gray area as to who must register (e.g. La Pira & Thomas 2013). A second problem is that registration by an interest group or lobbyist provides a right to lobby, but does not necessarily mean that the group has lobbied in the jurisdiction. In fact, it is not uncommon to find many registered interest groups with zero lobbying expenditures. Even when this issue does not occur, the groups that register to lobby are not a random sample of groups, as there is likely to be endogenous selection into registration, accompanied with all the aforementioned issues that arise with that. Finally, registries are composed of counts of interest groups, but do not provide information on lobbying effort by groups. This drawback will be particularly problematic if the goal of research is to identify how “active” or “effective” interest groups are in lobbying.

3. Transaction Reports

A few national governments and a significant number of state governments in the U.S. have moved beyond simply requiring lobbyists to register—and now collect transaction related data on lobbying. This data includes how much each interest group spends on lobbying in each
time period, how much (and from whom) each lobbyist earns in revenue in each time period, and, in some cases, in which general issue areas the interest group is lobbying. Advancements in the statistical analysis of lobbying have been driven, in part, by the availability of these new databases. One database, in particular, has been used extensively. The Lobbying Disclosure Act of 1995 (and its Amendments) requires lobbyists to report lobbying expenditures and other data to the U.S. Congress; this data has proven to be a popular source for papers on the subject. Papers using this federal data (Baumgartner et al. 2009; Bertrand et al. 2012; Blanes i Vidal et al. 2012; Richter et al. 2009) and similar state-level transactional data (Grasse & Heidbreder 2011; Lewis 2013; de Figueiredo & Cameron 2012; Cameron & de Figueiredo 2013; de Figueiredo 2014) now abound.

One advantage of this transactional data is its scope: researchers can now identify the timing, intensity, and focus on an interest groups’ lobbying effort. This should, in turn, allow researchers to link lobbying intensity to lobbying outcomes. Second, this data can be integrated into the registry data to develop better network maps of the lobbying process and help identify endogenous selection and timing issues in registries. Third, because of the quantity of the data in both time and transactions, this data supports “big data” studies with large numbers of observation units over long periods of time that will permit more advanced statistical methods to be employed.

Despite the attractiveness of transactional data, it is not without disadvantages. Like registry data, there is some gray area as to what is to be included in lobbying expenditures. Moreover, these rules and regulations differ across jurisdictions, making cross-jurisdiction comparisons difficult without cross-jurisdiction fixed-effects. Second, the transactional data show lobbying intensity, but does include the content of the message that is being transmitted.
Third, the US federal transaction data does not inform the researcher who in the legislature, for example, is the specific target of the lobbying effort. Thus, unlike campaign contributions, researchers cannot map interest groups’ lobbying efforts to individual legislators.

B. FUTURE PATHS FOR DATA COLLECTION ON LOBBYING

The current data now available for researchers in lobbying should help to advance the empirical research agenda in lobbying. In particular, the use of transactional data, which is now available over longer time periods, will likely be one of the more attractive archival data roads for researchers to travel in order to employ the more advanced statistical methods to obtain better statistical identification and isolate the causal effects of lobbying on policy.

However, the transactional data alone will likely not be enough. In order to make headway, researchers will have to integrate archival datasets (transactional data and registry data) with external datasets to obtain natural experiments and better statistical identification for isolating causal mechanisms. Blanes i Vidal et al. (2012) integrated a database on lobbyists’ career histories with transactional lobbying data; Kang (2013) integrated a database on bill proposals with transactional lobbying data; de Figueiredo & Silverman (2006) integrated a database on university overhead rates and academic earmarks with transactional lobbying data; Richter et al. (2009) integrated a database on corporate tax payments with transactional lobbying data. Integrating external archival datasets with the current lobbying datasets will likely yield substantial payoffs for research.

In addition, researchers will likely profit from expanding statistical research on interest group lobbying outside of the U.S. and its states. Papers have been written using datasets for a number of countries, including China (Yadav 2008), India (Yadav 2008), Japan (Naoi & Krauss...
2009), Mexico (Siegel 2005), and Norway (Alt et al. 1999). Although lobbying data in developing countries is scant, data collected in these countries will also allow researchers to make substantial contributions to understanding the breadth of applicability of theories of lobbying. Studies that employ this international data could contribute substantially to defining the generalizability of the U.S.-based empirical work. Even simple bivariate correlations in these less studied international contexts could be quite useful until higher quality data is available and more well-defined causal inference studies can be designed.

Finally, archival data will benefit from being combined with survey data, despite their many disadvantages, to uncover new ground and institutional details in the mechanisms that cause lobbying outcomes (e.g. Baumgartner et al. 2009).

V. FUTURE RESEARCH

In thinking about the future of empirical research in lobbying, there are a number of areas where crisp, clean, and well-identified statistical work could contribute substantially to our understanding of lobbying and help the field to make substantial headway in answering the questions posed below. In addition to simply applying quasi-experimental approaches, understanding and incorporating the institutional details of lobbying remains very important in the research design stage and will likely lead to the highest value results. Good empirical studies in the areas suggested below could substantially move the literature on lobbying.

The first area that deserves attention is, “Why is there so little money is lobbying?” In the United States, federal budgets represent in excess of two trillion dollars, yet lobbying represents only three to four billion dollars. Relatedly, “Why do so few interest groups lobby?” If lobbying is presumed to be so influential in appropriations and most policy domains, then
empirical researchers must convincingly guide scholars to an answer of why approximately only 10% of firms lobby in practice and why they spend so little given the magnitude of potential benefits politicians could redirect towards interest that do lobby.

Second, a related area that deserves substantially more attention is the quantification of the returns to lobbying. While lobbying is pervasive in the American and many other political systems, the returns to lobbying cannot be infinitely large. Moreover, the billions of dollars being spent annually on the activity are likely not all wasted. Presumably, the interest groups that choose to lobby have some expectation of a return; however, understanding more about the distribution of the payoffs among interest groups would seem to be an important question to answer. A more nuanced question might be, “When, and under what conditions, does lobbying produce a payoff?” If there are substantial marginal returns to lobbying, as most papers suggest, why is there not much more investment? Put another way, when do the marginal benefits of spending an extra dollar on lobbying begin to be outweighed by the marginal costs?

de Figueiredo & Silverman (2006) suggest that there are returns to lobbying only when the supply and demand conditions are in equilibrium. Moreover, the author’s argue because a legislator’s time is extremely limited and because the legislator is seeking a particular piece of information that will be persuasive, once the legislator has been persuaded, the marginal returns to additional lobbying is likely zero. An alternative view is that lobbying that the returns to lobbying is extremely difficult to measure because most lobbying is defensive, preserving the status quo. Measuring the impact of lobbying to get no measurable policy change makes it difficult for researchers to measure the returns to lobbying. In addition, if interest groups lobby friendly legislators who subsequently and privately lobby marginal legislators for votes (Hall & Deardoff 2006), quantifying the precise returns to lobbying is difficult. Employing better
datasets and more advanced empirical techniques, researchers should now be able to come closer to isolating and quantifying the effect of lobbying expenditures on policy outcomes.

Progress in these first two areas of academic inquiry should allow us to make headway in answering additional questions of public policy interest. To date, the data suggests that businesses are represented slightly more than non-business interests in lobbying, but businesses spend substantially more on lobbying. However, Baumgartner et al (2009) find that although citizen groups lobby less and on fewer issues than business groups, they are more likely to be considered an important actor in such advocacy efforts. Do asymmetries in numbers or spending make a difference in influence? Or do businesses have to spend more on lobbying to make their voices heard? Are the payoffs large to business interests for higher lobbying effort? Or is the marginal return to business interests per dollar spent comparable to, or less than that, for citizen interests? By answering the questions in the first two areas outlined above we will make headway on these latter questions.

The third area that deserves more attention is quantifying the importance of “who you know” (connections) and “what you know” (expertise) in lobbying. It seems likely that both play a role in determining the value of interest group lobbying (Bertrand et al. 2012), but understanding their relative importance in the lobbying process would be helpful. In particular, we now have evidence that “who you know” generates substantial revenue for lobbyists (Blanes i Vidal et al. 2012). However, we have no direct empirical tests of the value of “what you know.” That is, there have not yet been direct tests incorporating the content of the lobbying message provided to legislators.
One prevalent view is that campaign contributions provide access to politicians that are the targets of lobbying activity (Austen-Smith 1995). It could also be the case that “who you know” plays the same role as campaign contributions in providing access to politicians. Senators could rely on former staffers to screen potential lobbyists based on the content of the interest group’s information. Former staffers understand the preferences and utility function of the legislator well, and they understand her time constraints, so they might serve as useful gatekeepers of information for the politician. Once the interest group is before the senator, only “what you know” may matter for outcomes obtained. Empirical studies of this phenomenon would seem to be a promising agenda.

Fourth, it would be helpful to know how intensely different legislators are targeted. Surveys have allowed researchers to identify who is targeted for lobbying by interest groups, but little research has been done on how intensely they are targeted and what kinds of messages are the most influential in a legislator’s decision. Further empirical work on this topic with more sophisticated datasets is likely to yield substantial fruit.

Fifth, in a broader study of political influence, scholars would benefit from understanding how interest groups allocate resources across different instruments (lobbying, campaign contributions, grassroots organization, endorsements, media campaigns, etc.) and which types of interest group pressure are most effective. This is a difficult question to answer because unobserved factors that influence lobbying may also affect other instruments (such as campaign contributions). Thus, very careful implementation of the quasi-experimental methods noted earlier would seem to be essential to making any credible headway in this area.
A final and very important avenue for researchers to pursue is the empirical testing of the implications of theoretical models (EITM). There are literally hundreds of theoretical models of lobbying and influence, with many different and frequently opposing predictions. Careful testing of the implications of some of these theoretical models where appropriate (as outlined in Clarke & Primo 2012), has largely eluded researchers to date. Austen-Smith & Wright (1996), de Figueiredo & Cameron (2012), Kang (2013), and Cameron & de Figueiredo (2013), all cited earlier in this essay, provide examples of empirical papers tightly and clearly linking theory and testing in a way that can be falsified. These papers help to support and reject broad classes of theoretical models. More work along this vein of research will allow the field to cull the vast theoretical literature into a set of core theoretical models that are most useful in explaining the actual practice of lobbying.

The empirical research on lobbying has progressed substantially over the past decade. New datasets, new methodologies, and new research designs together create the opportunity to not only investigate some of the more tenuous results currently found in the literature, but also to answer some core research questions. Employing these new techniques will advance empirical research in lobbying substantially over the next decade.
REFERENCES


