The Political Economy of Publicly Funded Biomedical Research: Evidence from NIH Funding for Rare Diseases

Deepak Hegde[1] and Bhaven Sampat[2]

July 2011

Preliminary: Please do not cite without permission

ABSTRACT

How do the scientific community, politicians, and special interest groups influence the allocation of publicly funded research? We investigate this question by analyzing novel data on the lobbying expenditures of disease interest groups, Congressional earmarks, and funding by the National Institutes of Health (NIH) for 955 rare diseases, during the 1998-2008 period. We show that Congressional earmarking for the rare diseases is responsive to the lobbying expenditures of interest groups associated with the diseases, but not to our measures of scientific opportunity and disease burden. NIH funding for the diseases is not responsive to earmarking overall, but only to earmarking associated with lobbying expenditures. We also provide suggestive evidence that lobbying has an informational role, helping focus Congressional attention on diseases with higher burden or scientific opportunity. Our findings have implications not only for science policy, but also for the political economy literature on how interest groups influence policy.

[1] New York University; email: dhegde@stern.nyu.edu
[2] Columbia University; email: bns3@columbia.edu

1 Introduction

Rett Syndrome is a developmental disorder, affecting about one in ten thousand children. The syndrome is primarily seen in females six-to-eighteen months old, and causes a loss of communication skills and of the purposeful use of hands. In severe cases it causes disorganized breathing patterns and seizures, leading to death. The syndrome has no known cures. In 1999, a discovery linking the disorder to mutations of a gene called MECP2 opened new avenues of research and potential treatment. In 2002, The International Rett Syndrome Association mobilized parents, friends of patients (including actress Julia Roberts), and scientists to lobby the U.S. Congress to increase the National Institutes of Health's (NIH's) funding for research on the disease. In response, the House Appropriations Committee included language in the reports accompanying its appropriations for the NIH, "encouraging research" on Rett syndrome. Figure 1a displays the relevant excerpt from the Congressional report. As the accompanying chart, Figure 1b, shows, in the year that followed, NIH grants for research on Rett Syndrome increased by 65%, from $4.6 million in 2002 to $7.6 million in 2003.

Figure 1 here

The National Institutes of Health is the world's largest single source of funding for biomedical research. Each year, the NIH allocates public funds appropriated by the Congress ($31.2 billion in 2010) to support research across hundreds of diseases and scientific fields. The agency allocates funds primarily through the peer review process, in which independent outside scientists are tapped to evaluate the scientific merit of research proposals submitted by investigators. In contrast to other U.S. R&D funding agencies (e.g. the Department of Agriculture and the Department of Defense), Congress generally does not include "hard" earmarks that set aside funds for specific projects and institutions in its appropriations for the NIH. Instead, as in the case of Rett Syndrome above, the Congress relies on what are frequently referred to as "soft" earmarks, through language that "urges" and "encourages" the NIH to support specific areas of research. These soft earmarks are contained in the text of the Congressional Committee reports that accompany NIH appropriations legislation (Hegde & Mowery 2008; Hegde 2009; Sampat 2009. Policymakers and scientists have expressed concerns that soft earmarks reflect the influence of powerful lobbying groups, and may distort allocations towards fields backed by powerful disease groups, and away from the best science (e.g. IOM 1998).

Does Congressional support for research across different diseases, in the form of soft earmarks, respond to the lobbying efforts of interest groups? How do bureaucratic agencies such as the NIH respond to Congressional directives? Do organized lobbying groups provide useful information to decision makers (e.g. Grossman and Helpman 2001), or do they seek rents at the expense of the larger public (e.g. Stigler 1971, Peltzman 1989)? Though anecdotal responses to these questions abound, systematic evidence on the mechanisms and effects of lobbying on the allocation of public funds are limited.

In this paper, we attempt to fill this void, examining NIH funding for research on 955 rare diseases during the 1998-2008 period. We focus on rare diseases for several reasons.[3] First, our analysis links

[3] A rare disease is generally considered to have a prevalence of fewer than 200,000 affected individuals in the United States. Rare diseases are also sometimes referred to as orphan diseases or neglected diseases.

data from disparate sources, and doing so reliably requires focus on a specific set of diseases. Rare diseases tend to be well-defined, discrete conditions (unlike more prevalent causes of deaths such as heart or respiratory diseases) and can be precisely mapped to the funding, earmarking, disease burden and publication data used in our analyses.[4] Second, although individual rare diseases have low prevalence, these diseases are collectively responsible for the deaths of hundreds of thousands of individuals annually. Third, private sector firms find it unprofitable to invest in R&D for diseases that are prevalent in small numbers; public funding is thus considered particularly important to find treatments for rare diseases (IOM 1998; Ashbury 1985).[5]

Our results suggest that a doubling of lobbying expenditures by interest groups associated with a particular disease is associated with a statistically significant 5 percent increase in the number of Congressional earmarks for that disease, *ceteris paribus*. (Here, and below we use the term "earmark" to refer to soft earmarks, except when otherwise specified). Our main models include disease-fixed effects, ameliorating concerns about threats to identification due to unobserved disease-specific factors. We also examine a quasi-experiment that involves a shock to the lobbying industry, brought about by changes in party control of the President's office and the Senate, and arguably unrelated to changes in disease-specific characteristics. We show that the estimated elasticity of earmarks with respect to lobbying responds to this change in a way that is consistent with a causal interpretation of our results on the effects of lobbying.

In addition, we find that earmarking is not significantly related to within-disease variation in our measures of disease burden and scientific opportunity. However, increases in disease burden and scientific opportunity, when accompanied by increases lobbying, do increase Congressional earmarks for the diseases. Even more striking, absent changes in disease burden or scientific opportunity, the estimated elasticity of earmarking with respect to lobbying is not statistically different from zero. We tentatively interpret this as evidence that Congressional earmarking responds to lobbying activity only when this activity provides information about changes in scientific opportunity and/or burden associated with diseases.

Do Congressional earmarks influence agency allocations for diseases? Here, we find that NIH funding for new projects does not respond to earmarks, overall, or directly to lobbying activity. However, two-stage least squares regressions that examine the variation in earmarks related to lobbying suggest that a 1 percent increase in lobbying-induced earmarks is associated with a 2.2 percent (at $p<0.05$) increase in NIH funding for the disease. We also find NIH funding through grant mechanisms that solicit proposals for research in specific diseases ("Program Announcements" and "Request for Applications") are particularly responsive to earmarking overall, and to earmarking associated with lobbying.

In addition to their implications for science policy, our analyses make several contributions to the interest group politics literature. First, our research addresses the question of how interest groups

[4] By contrast, estimating the effects of lobbying on earmarks and funding for cancer would require assigning lobbying, earmarks, and funding for "cancer" to specific types of cancer.

[5] The 1983 Orphan Drug Act aimed to create tax and market exclusivity incentives for stimulating private sector research on rare diseases (Ashbury 1985).

influence Congressional priorities which in turn influences policy choices (Grossman and Helpman 2001; Baumgartner and Leech 1998; Jones and Baumgartner 2005). While the existing literature emphasizes the close interplay among interest groups, politicians and federal bureaucrats in shaping policy outcomes, few studies provide evidence on the links connecting these three vertices of the "iron triangle" of American policymaking.[6] The limited previous evidence on these issues reflects several challenges. First, the pathways through which interest groups affect politicians and politicians influence agencies typically do not have clear footprints. Second, counterfactual allocations against which to assess the effects of political influence typically do not exist. In addition to being important, the NIH allocation process which we study is data rich: we can link disease interest groups to their lobbying expenditures; we can control (albeit imperfectly) for the demand-side characteristics (e.g. disease burden) as well as scientific opportunity which influence counterfactual allocations; and, most importantly, we can link Congressional influence (earmarks) and agency allocations to specific diseases.

Second, we examine the effects of a novel mechanism through which politics can influence agency allocations: "soft" earmarks. Much of the literature on lobbying and earmarks (e.g. de Figueiredo and Silverman 2006; Evans 2006) focuses on "hard" earmarks. Although soft earmarks in NIH appropriations bills have generated controversy, there is no evidence they affect NIH funding choices. Deepening our understanding of soft earmarking and its effects also may have important implications beyond the the NIH context. Ethics rules passed in 2007 to curb hard earmarks in the U.S. budget more generally may have shifted interest group strategies towards soft earmarks, which are not subject to the same scrutiny or restrictions (Nixon 2010).[7] Yet there is no empirical evidence on whether and when such soft-earmarks – which, unlike hard earmarks, are loosely specified and don't indicate dollar amounts--actually influence agencies' funding choices. Here, we provide the first evidence that they do.

Finally, a perennial debate in the history of the NIH is the role that political influence has, and should have, in its allocation processes.[8] Some legislators and disease advocates have argued that since the NIH's peer review process primarily responds to scientific opportunity, political influence is required to introduce the public's health considerations and priorities into the process. On the other hand, scientists involved in the allocation process generally resist political influence, alleging it provides a channel for powerful interest groups to distort allocations in their favor. Our analysis suggests that the "politics versus science versus health" distinctions may be too simplistic: the three sources of influence interact in interesting ways.

[6] The overwhelming majority of studies that consider the influence of the iron triangle on policy making tend to be descriptive. A recent exception is Iaryczower, Spiller, and Tommasi (2004) which links the theory of interest group influence over the legislature with that of congressional control over the judiciary and tests the theory in the context of Supreme Court labor decisions in Argentina

[7] A Congressional Research Service (CRS) report notes "A soft earmark, on the other hand, is an expression using terms such as should, urges, endorses, or recommends. Since both hard and soft earmarks are an expression of congressional intent, some argue that there is little if any distinction between the two, and the Administration must treat each with equal weight."

[8] This debate reflects broader tensions in the history of science and technology policy between the scientific community and politicians, dating back to the Bush-Kilgore debates at the end of the Second World War about how public science should be governed.

The remainder of this paper proceeds as follows. Section 2 discusses the institutions involved in the allocation of funds for biomedical research. Section 3 specifies the empirical model we estimate, and discusses the data. Section 4 reports estimates of the effects of lobbying on earmarking for rare diseases. Section 5 investigates the influence of Congressional earmarking on NIH allocations. Section 6 concludes.

2 Institutional Setting

2.1 The National Institutes of Health and peer review

The National Institutes of Health is part of the U.S. Department of Health and Human Services and provides 85 percent of total federal support for R&D in the biological, medical, and psychological sciences (based on FY2008 federal obligations, NSF 2010). More than 80 percent of the agency's funding is awarded annually through competitive grants to researchers at over 3,000 universities, medical schools, and other research institutions.[9] The agency is organized into 27 independent Institutes and Centers which specialize by disease (e.g. National Cancer Institute), organ (National Eye Institute), field of science and medicine (e.g. National Institute of General Medical Sciences), or by stages of human development (e.g. National Institute on Aging) (McGeary & Smith, 2002).

Figure 2 here

Figure 2 plots annual extramural allocations by the NIH for research in rare diseases, and all other funding, for fiscal years 1998-2008. The sharp increases of funds during the 1998-2003 period (from $1.95 billion to $3.12 billion for rare diseases, and from $14.5 billion to $23.8 billion for other research) reflects the success of an unprecedented bipartisan effort to double the NIH's budget during this five-year period.

The individual Institutes at the NIH utilize a "dual peer review" process to evaluate proposals from researchers. In the first stage of this process, grant applications are evaluated by panels of external scientists from the relevant fields. These experts score applications based on their significance, technical merit, innovativeness, and investigators' qualifications. Acceptable applications are assigned to the NIH Institute or Center best suited to fund the research where they are again reviewed by a "National Advisory Council" composed of scientists and public representatives. Each Institute's (or Center's) Advisory Council recommends applications for funding by considering priority scores and the proposed project's relevance to the Institute's mission. The Director of the Institute/Center makes the final funding decision based on the relevant Advisory Council's recommendation (NIH 2008).

The peer review system at the NIH is considered by many experts to be insulated from political influence. Drazen and Ingelfinger (2003) note that "the selection of research for funding [for investigator-initiated grants] is based solely on merit; politics has no place in this system" (2259). The former Director of the

[9] The rest of the funds support "intramural" research, or activities conducted scientists working at NIH's campus in Bethesda.

NIH, Harold Varmus, suggests Congressional reluctance to include hard earmarks in the NIH budget "represent[s] votes of congressional confidence in the NIH's system of peer review" (Varmus 2010, 150).

2.2 The Congressional appropriations process and earmarks

Annual Congressional decisions concerning the NIH budget are made by the Labor, Health and Human Services, and Education and Related Agencies Subcommittee (LHHE) of the House and Senate Appropriations Committees. The bills reported out to the floor of the House and Senate by each chamber's appropriations committee each year indicate the total budgets for each of the NIH Institutes and Centers every year, one means by which Congress can shape NIH allocation patterns.

Soft earmarking, through the report language recorded in the annual appropriations committee's meeting reports, is another way. For example, House Report 109–515 accompanying the appropriations bill for Fiscal Year 2007 includes earmarks for Interstitial Cystitis ("research on IC is still in its infancy ... the Committee encourages NIDDK [the National Institute of Diabetes and Digestive and Kidney Diseases] to place emphasis on Interstitial Cystitis-specific funding in order to focus on the basic science of IC and to attract and sustain research in the field") and for stroke research ("the Committee continues to place a high priority on stroke research and encourages NINDS [the National Institute of Neurological Disorders and Stroke] to al locate resources to basic, clinical and translational research into stroke.") The analogous Senate report also includes scores of directives, including for Marfan Syndrome ("the Committee urges the Institute to continue to support the major advances made in this area through al l available mechanisms") and Batten Disease ("the Committee strongly urges the Institute to increase funding for Batten disease research by actively soliciting grant applications and taking aggressive steps to assure that a vigorous research program is established.")

As noted earlier, these soft earmarks have been the subject of considerable controversy in policy discussions about how the NIH does and should make allocation choices. Some argue that earmarks are the only channel through which considerations of "public needs" can be injected into the allocation process, which (they contend) is otherwise narrowly oriented towards scientific opportunity. One advocate (quoted in Dresser 2001, 80) notes that "some patient groups believe the decision- making process at the NIH is basically a closed process where patient organizations are only consulted at later stages, when decisions in fact have already been made." However, others worry that earmarks target diseases without concern for the scientific feasibility of research, and that this low quality research crowds–out funds for higher quality peer reviewed research (Greenberg 1998). Representative Greg Ganske (R–IA) has argued that earmarks "are turning the floor of Congress into a scientific peer review panel" and questioned "whether members of Congress have the scientific expertise to determine where the most promising areas of research are" (IOM 1998).

Another concern is that these earmarks may cater to powerful disease interest groups rather than the public more generally. One potentially problematic aspect of earmarking is that certain diseases may be more "politically correct": for example, some are better able to generate media attention, perhaps with the help of celebrity spokespeople (cf. Armstrong et al. 2006). A related concern is that some disease communities may be inherently more difficult to organize. For example, diseases with few survivors (e.g.,

lung cancer or pancreatic cancer) or those which are associated with a social stigma (e.g., depression or urinary incontinence) may have few advocates. In other words, while earmarks may be justified as a means to inject health considerations into NIH allocation decisions, if there is differential ability to organize across diseases, there is little reason to believe that such a system would result in a just or efficient allocation of resources (Dresser 2001).

Concern that earmarks reflect disease group lobbying activities and distort NIH allocations fueled considerable controversy during the 1990s, including Congressional hearings with testimony by the NIH Director (Varmus 1997) and an important Institute of Medicine inquiry into NIH priority setting (IOM 1998).

2.3 Disease interest groups

Lobbying for earmarks represents a new chapter in the long saga of interaction between disease groups, Congress, and the NIH. Throughout the agency's history, advocates and interest groups have played an important role in mobilizing public support and Congressional influence for research funding (Drew 1967, Cook-Deegan and McGeary 2006). For example, disease groups were important in generating political and public support for medical research after World War II (Strickland, 1972; Drew 1967). In recent decades however, interaction between advocates and the agency grew more contentious, as research advocacy became more disease-specific. Thus in the 1990s, powerful and well organized advocates for specific diseases--AIDS and breast cancer in particular--began to lobby for increased research on these diseases, including through obtaining Congressional report language to bypass the peer review process. These activities fueled the concerns, discussed above, that powerful groups, rather than actual disease burden was increasingly driving NIH allocations. A 1997 Congressional Research Service Report noted:

> "In this atmosphere of impeding budgetary constraint, health advocacy groups find themselves increasingly at odds with one another, lobbying congressional offices and NIH for more research on their specific disease of interest rather than for health research in general. Such lobbying efforts appear to have succeeded in gaining large increases for certain diseases (e.g., AIDS and breast cancer) at the expense of others. When budget resources are limited and not growing, adding funds to one area almost inevitably limits funds to another. This more than likely has added to the intensity of already fierce lobbying for disease specific earmarks in the NIH budget."

Indeed, concern that a "large and growing set of groups pushing for earmarks" (IOM, 1998) was interfering with NIH priority setting was one of the motivations for the IOM report cited above.[10]

[10] A 1996 Congressional Research Service Report observes: "Many representatives of disease advocacy groups claim that in the past they pushed solely for increased overall funding for basic research at NIH. However, after years of perceived neglect they are now intent on following the example of the AIDS and breast cancer lobbyists and are promoting increases for their area of interest alone. These groups believe increased lobbying of Congress and NIH is the only way they will receive more equitable funding and attention for their cause" (CRS 1996).

Disease interest groups seek to influence policy though a number of channels: by directly lobbying Congress, organizing testimony, as well as mobilizing grassroots advocacy campaigns, e.g. encouraging patient groups to contact their representatives. Expenditures on the first two activities are counted as "lobbying" expenditures in the empirical analyses below.

3 Empirical specification and data

3.1 Empirical model

We begin with a stylized model of optimal NIH funding, adapted from Lichtenberg (1998). The model is simplistic and ignores the complexity of the allocations process, but helps provide structure to our empirical analyses. It is similar in spirit to previous less formal characterizations of optimal allocation rules for biomedical research (e.g. Zeckhauser 1967; Weisbrod 1983).

Following Lichtenberg, we start by assuming the presence of two diseases. Let M_i be the number of people killed by the disease i (i=1, 2).[11] Let π_i be the probability of finding a treatment for the disease i. π_i is a concave function of research funding for the disease, Y_i, and a disease-specific scientific opportunity parameter, P_i such that $\pi_i = P_i Y_i^{\alpha}$ where $0 < \alpha < 1$. We assume that the total research budget, Y, is exogenously given and fixed, such that $Y = Y_1 + Y_2$

In this model, policymakers attempting to maximize the expected total number of people cured of both diseases would choose Y_1 to maximize:

$$W^* = M_1\pi_1 + M_2\pi_2 = M_1 P_1 Y_1^{\alpha} + M_2 P_2 Y_2^{\alpha}$$

$$= M_1 P_1 Y_1^{\alpha} + M_2 P_2 (Y - Y_1)^{\alpha} \quad (1)$$

The first-order condition implies that relative funding of research on the two diseases should satisfy:

$$\ln\left(\frac{Y_1}{Y_2}\right) = \left(\frac{1}{1-\alpha}\right)\ln\left(\frac{M_1}{M_2}\right) + \left(\frac{1}{1-\alpha}\right)\ln\left(\frac{P_1}{P_2}\right) \quad (2)$$

Writing $\beta = \left(\frac{1}{1-\alpha}\right)$ and generalizing the model to the case of i>2 diseases, we obtain i-1 equilibrium conditions of the form:

$$\ln Y_i = \text{constant} + \beta \ln M_i + \beta \ln P_i \quad (3)$$

Equation 3 implies that, in this model, public funding (Y) for a given disease (i) should be an increasing function of disease burden and scientific opportunity.[12]

[11] In the analyses below, we use the number of deaths associated with a disease as an indicator of disease burden. However, other measures (e.g. costs of a disease, incidence, disability-adjusted life years) could also be used.

[12] This is consistent with the views of various experts and scholars on the desired determinants of applications. For example, the aforementioned 1998 IOM Report on the structure of NIH allocations is titled: *Scientific Opportunities*

With cross-sectional data on measures of disease burden and scientific opportunity, we could estimate β. In the above model, scientific opportunity enters into the objective function the same way as disease burden, hence the same parameter β for both. However, we do not restrict actual allocations to be identically responsive to both disease burden and scientific opportunity and separately estimate β_1, which captures the effect of M, and β_2 which captures the effect of P, on allocations (as in equation 4). Since M and P enter the expenditure function as log terms, the estimated coefficients have elasticity interpretations. Thus, we can write the equation to be estimated as:

$$\ln Y_i = \beta_0 + \beta_1 \ln M_i + \beta_2 \ln P_i + e_i \quad (4)$$

A key empirical challenge for isolating the influences of the two disease characteristics (M and P) on allocations is controlling for other attributes of the diseases, e.g. prevalence, the income levels of the populations affected by the diseases, or private funding for research. These other attributes of the diseases may be correlated with *M*, *P* and *Y*, and if omitted from the OLS regressions, lead to biased estimates. Since it not feasible to observe and measure all these "other" attributes of the diseases, we use disease-specific fixed effects to control for such confounders. These disease-fixed effects ($\sum I_i$) account for unobserved disease-specific influences on funding that remain constant across the years of the study.[13] Hence, with panel data, we can hold constant disease-level characteristics ($\sum I_i$ representing a vector of disease-specific indicator variables) and estimate how within-disease funding support (Y) responds to changes in M and P. The corresponding equation can be written as:

$$\ln Y_{i,t} = \beta_0 + \beta_1 \ln M_{i,t} + \beta_2 \ln P_{i,t} + \sum I_i + e_{i,t} \quad (5)$$

Our analysis spans 11 years, a relatively short period of time. With disease fixed effects, there may not be sufficient within-disease variation to estimate the effects of disease burden or scientific opportunity. However, the primary goals of our analysis are to understand: (i) the relationship between the lobbying efforts of interest groups and Congressional earmarking for diseases and (ii) the relationship between Congressional earmarking and NIH allocations, controlling for factors that influence allocations and political support. Accordingly, we extend Equation 5 to examine how disease-level support responds to additional factors such as the lobbying efforts of interest groups (L):

$$\ln Y_{i,t} = \beta_0 + \beta_1 \ln M_{i,t} + \beta_2 \ln P_{i,t} + \beta_3 \ln L_{i,t} + \sum I_i + e_{i,t} \quad (6)$$

In estimating the model, we first focus on the effects of lobbying expenditures (L) of interest groups on Congressional support for diseases (Y^C) controlling for burden (M) and productivity (P). Funding levels during the past year may influence both lobbying intensity and Congressional earmarking. Accordingly, we include NIH funding in the previous year (*K*) as an explanatory variable. Year-specific indicator variables ($\sum T_t$) control for time-trends in funding. Y^C, *P*, *M* and *L* each vary across the diseases (*i*) and

---

and Public Needs. The IOM Report also emphasizes other dimension that this model does not, including equity considerations, effects on costs, and extent of market failure (cf. Garber and Romer 1996).

[13] If i the disease-specific unobserved attributes were constant during the study period, the fixed-effects model is the statistical equivalent of a controlled experiment in which the only change for a given disease is the change in the observable variable of interest. The difference in the outcome (Congressional support) before and after this change could then be interpreted as the causal impact of the observable variables of interest.

over time (*t*); the unit of analysis in our study is disease-year. Thus, the equation that estimates lobbying influence on Congressional earmarking is:

$$\ln Y^{C}_{i,t} = \beta_0 + \beta_1 \ln M_{i,\,t-1} + \beta_2 \ln P_{i,t-1} + \beta_3 \ln L_{i,t-1} + \beta_3 \ln K_{i,t-1} + \sum I_i + + \sum T_t + e_{i,t} \qquad (7)$$

Next, we examine NIH's grants for specific diseases (Y^{NIH}) as a function of *M* and *P* for each disease-year. Since we expect lobbying to influence NIH's funding not directly, but through their effect on Congressional support (Y^{C}), we use an instrumental variables framework to estimate this indirect effect. Specifically, we use the effect of lobbying on earmarks obtained by estimating (6) in the first-stage to investigate the effect of (lobbying-driven) Congressional earmarking on NIH's funding in a second-stage. Thus, we examine the effect of earmarks and other disease-specific characteristics on NIH's allocations by estimating the following two-stage model, with L as an instrument for Y^{C} :

$$\ln Y^{NIH}_{i,t} = \beta_0 + \beta_1 \ln M_{i,\,t-1} + \beta_2 \ln P_{i,t-1} + \beta_3 \ln Y^{C}_{i,t-1} + \beta_3 \ln K_{i,t-1} + \sum I_i + + \sum T_t + e_{i,t} \qquad (8)$$

3.2 Data and descriptive statistics

Rare diseases have no official definition, but are typically characterized as those with prevalence of less than 200,000 individuals. The National Organization for Rare Diseases (NORD) identifies 1,200 such diseases in its Rare Disease Database. Since some of these are not "true" rare diseases (e.g. HIV-AIDS) we focus on the subset of 955 that were also present in the NIH's own rare-disease database.[14] Even this list includes some diseases that have high prevalence. Thus the NIH cautions that the list "should not be used as a reference or guarantee that a condition is rare" and that "certain diseases with 200,000 or more affected individuals may be included in this list if certain subpopulations of people who have the disease are equal to the prevalence standard for rare diseases." Accordingly, to be cautious, the use of the term "rare disease" in this paper should be interpreted as "a disease in both the NORD and NIH rare disease databases," a set that may include some diseases with relatively high prevalence.

The NORD database also provides all known synonyms for each disease, which is indispensible for collecting data on the characteristics of the diseases from five different sources as described below.[15] The NORD data also include information on all organizations associated with each of the diseases, allowing us to link them to lobbying data.

3.2.1 Dependent variables

(i) Congressional support for diseases (Y^{C}), measured by the number of "soft" earmarks

We collected data on all report language that mentioned support for the 955 diseases (and/or their synonyms) from the House and Senate Labor/HHS/Education subcommittees' appropriations reports accompanying appropriations bills for FY1998 through FY2008. We read each of these reports for each year and chamber, and hand-coded as an earmark any mention of a rare disease in conjunction with words indicating support (mainly "urge," "encourage," and "direct"). During our study period, House

[14] http://rarediseases.info.nih.gov/RareDiseaseList.aspx?PageID=1

[15] On average, each of the 955 diseases has 3.2 synonyms; the median number of synonyms is 2.

reports specified an average of 56 earmarks for rare diseases, with some diseases supported by more than one earmark in the same report. Senate reports specified a higher number of earmarks on average –85 per report – during the period. In an average year, the Congressional reports specify 142 earmarks, supporting 32 rare diseases. The discrepancy between these numbers reflects that some diseases are the subject of earmarks in appropriations bills for multiple NIH Institutes. In such cases we count each earmark as a separate instance.

Interestingly, rare diseases receive a disproportionate number of earmarks. For example in year 2006, House and Senate reports together specified a total of 310 earmarks of which 140 (or 45% of earmarks) focused on rare diseases. By comparison, Figure 3 shows that the 955 diseases in our data received around 15 percent of NIH's overall funds, and account for roughly the same percentage of deaths in the U.S., and of NIH-funded publications indexed by MEDLINE:

Figure 3 here

(ii) Agency allocations for diseases (Y^{NIH}), measured by NIH's peer-reviewed funds

Our primary source of data on NIH funding is the CRISP (Computerized Retrieval of Information on Scientific Projects) database. The CRISP data includes titles, abstracts, principal investigators, institutions, funding institute, and amounts for all of NIH's extramural grants and contracts since 1972. Though previous work (Toole 2007; Lichtenberg 1998) has used "thesaurus" terms in the CRISP data as a measure of funds by disease, we found the thesaurus terms unreliable for rare diseases. For example, many grants referenced a rare disease in the title and abstract, but not in the thesaurus keywords. And many rare diseases did not have entries in the CRISP thesaurus at all.

Instead of using CRISP keywords, we searched for the names of the 955 diseases (and their synonyms) in the abstracts of each of the 600,000 grants awarded between 1998 and 2008, associating a grant with a disease if any of the words in the abstract matched one of the 955 rare diseases or their synonyms. Overall, the 955 diseases mapped to 77,005 distinct grants during the period. We used a separate NIH database, RePORTER, to collect information on the amount of funds associated with each grant, and aggregated the funds received by each disease for each year. In an average year during 1998-2008, the 955 diseases in our set received 7,000 grants, or $2.8 billion in total funding.

How well does our search do? Historically, the NIH has been reluctant to provide data on funding by disease (Sampat 2011). This changed with the Research and Conditions Disease Categorization (RCDC) initiative launched in 2008, but these data are available too late for our purposes, and focus only on a set of diseases of interest to Congress. For the 33 diseases in our sample also listed on the RCDC, we compared our funding numbers to those from the RCDC. The correlation between the two is .92, suggesting the search strategy described above generates reliable figures of funding by disease, at least if the RCDC is treated as the gold standard.

3.2.2 Independent variables

(i) Disease burden (M), measured by deaths caused by diseases

Disease burden associated with diseases is multi-faceted, and the "right" measure of disease burden on society is a matter of considerable controversy and debate (Gold et al 2002). Disease incidence, costs, and quality-adjusted life years lost to a disease are common measures. Unfortunately we were unable to collect data on numerous dimensions of disease burden for the 955 diseases, in part because standard data sources to construct these measures (e.g., the Medical Expenditure Panel Survey) are based on surveys of the population, and lack the statistical power to generate reliable population-level estimates for rare diseases (which, by definition, are of low prevalence).

Instead, we use data on one measure that is not only reliable, but also collected annually: the number of deaths associated with each disease, as reported in the Multiple Cause of Death Mortality files, developed through the National Vital Statistics System of the National Center for Health Statistics.[16] These data are a census of all deaths in the U.S., and are listed by ICD-10 code. We determined the ICD-10 code for each disease (and/or synonym) in our sample, and used this correspondence to construct data on deaths by disease and year. The 955 diseases in our data accounted for between 13 and 15 percent of all deaths (attributable to all diseases listed in the CDC data) each year during the span of our study. The 955 diseases collectively are responsible for 307,182 U.S. deaths during an average year of our study. Column 1 of Table 1 reports trends in the deaths associated with these diseases.

Table 1 here

(ii) Scientific opportunity (P), measured by number of scientific publications in disease-field

Scientific opportunity is difficult to measure across diseases; indeed, NIH peer reviewers spend hours discussing it for individual applications. The measure for scientific opportunity we use relies on information from MEDLINE on publications associated with the diseases. We mapped each disease (and its synonyms) to a MeSH (Medical Subject Heading) entry in MEDLINE, and constructed disease specific publication stocks by year. The idea behind this measure is that disease areas where there is a change in scientific opportunity should also see a change in publications. While admittedly imperfect--for example, publications could reflect scientific fads rather than objective opportunity--we believe this measure does capture what NIH reviewers consider when evaluating applications: evidence that research in an area will yield results. On average, there were 19,836 rare-disease related publications per year, roughly 15% of all MEDLINE publications related to diseases.

(iii) Lobbying intensity (L) measured by lobbying expenditures of disease interest groups

The Lobbying Disclosure Act of 1995 requires all organizations that lobby Congress to disclose to the Secretary of the Senate's Office of Public Records (SOPR) good-faith estimates of all lobbying-related expenditures. Under this Act, lobbying expenditures are defined to include both in-house lobbying and hiring of external lobbyists. The Center for Responsive Politics (CRP) has collected these data for each organization that lobbied Congress and federal agencies in each year since 1998.[17] We identified in the

[16] We are currently supplementing these data with a nationwide census of all hospitalizations for each of the diseases—a cost-weighted measure of utilization—from the Health Care Utilization Project (HCUP).

[17] An organization that spends less than $10,000 in any six-month period does not have to state its expenditures. In those cases, the Center treats the figure as zero http://www.opensecrets.org/lobby/methodology.php

CRP data any organizations associated with the 955 diseases in our data (as indicated in the NORD organizations data). Using this information, we constructed annual amounts of lobbying associated with each of the diseases, from 1998 to 2008. During this period, 98 unique organizations reported positive lobbying expenditures. Some of these organizations are associated with lobbying for multiple diseases. In these cases, we divided the total lobbying expenditures reported by the organization by the number of diseases associated with the organization.

Figure 4 here

We emphasize that we do not know the extent to which lobbying expenditures are directly for earmarks for research funding, versus other activities (e.g., changing Medicare reimbursement rules). Our informal discussions with disease advocates for rare disease groups, and a review of their websites, suggest the bulk of their lobbying activities are oriented towards obtaining research funding. However, to be conservative, in the empirical analyses below we interpret lobbying expenditures as a proxy for the intensity of advocacy for a rare disease at a point in time, and do not attempt to directly estimate financial "returns" to lobbying (as in de Figueiredo and Silverman 2006).

Figure 4 shows trends in the overall lobbying expenditures for the diseases in our sample. The groups spent $12 million on lobbying activities in 2008.[18]

Table 2 here

Table 2 describes the five variables discussed above, for the 10,505 disease-year observations (955 diseases, 11 years) used in our empirical analyses. Notice that the average number of yearly earmarks is small, since only a small fraction of the 955 diseases--on average 36 rare diseases--have any earmarks in a given year.

4 Congressional support for rare diseases

4.1 Relationship between lobbying and soft earmarks

The first three columns of Table 3 report ordinary least squares (OLS) estimates of equations (4), (5), and (7).[19] Column 1 shows that a 1 percent increase in past-year deaths and publications are each associated with a 0.02 percent increase in Congressional earmarks for the disease (each is statistically significant at $p<0.01$). This baseline model with measures of burden and opportunity explains about 7 percent of the variation in congressional support across rare-disease years.

Table 3 here

[18] By comparison, the Center for Responsive Politics reports that the pharmaceuticals and health sectors overall spent $240 million on lobbying activities in that year.

[19] Equations 4 and 5 are estimated with the explanatory variables lagged by an year to account for the time-structure of the relationships between disease-characteristics and Congressional support

Column 2 shows that the estimated effects of deaths and publications are weaker when disease-specific fixed effects are included, although the estimated elasticity of publications remains statistically significant (estimated β_2 = 0.01; significant at p<0.05). The inclusion of disease-fixed effects also increases the amount of explained variation to 66 percent. Column 3 introduces lagged lobbying expenditures (L), our main variable of interest. A 1 percent increase in lobbying is associated with a 0.04 percent increase in earmarks (significant at p<0.01).[20] This estimated effect, which controls for unobserved heterogeneity among diseases with fixed effects, is nearly four times the estimated effect of changes in publications on earmarks.

In the models discussed above, the dependent variable is the natural log of earmarks. We also examined the effects of lobbying on the number of earmarks (in Column 4), and on the probability that a disease receives an earmark at all (in Column 5). A 1 percent increase in lobbying in the previous year is associated with 0.20 more earmarks (recall that the mean disease-year in the data receives 0.09 earmarks) and a 0.02 percent increase in the probability that the disease receives one. Lagged funding is not significantly related to earmarking in any of the specifications.

4.2 Omitted variable bias?

Overall, across the models, Congressional earmarking appears responsive to lobbying. A concern, however, is that the estimated effects may reflect the influence of unobserved (time-varying) disease-specific factors driving both lobbying and earmarking. For example, it is possible that unmeasured changes in disease burden or scientific opportunity (correlated with both lobbying and earmarks) are responsible for our estimated relationships, rather than the causal effect of lobbying on Congressional support for the diseases. This possibility that our estimates are biased cannot be ruled out absent an exogenous shock to lobbying. We attempt to mitigate this concern through a quasi-experiment, by investigating potential variation in Congressional response to lobbying brought about by changes in the political structure and the lobbying industry in Washington.

In 2001, Republicans gained control of the Presidency and Democrats of the Senate. (Republicans continued their control of the House gained in 1995.) Following this change, the lobbying industry experienced a shakeout: several lobbyists went out of business and others took up jobs as Congressional staffers in the new administration (see Kaiser 2009, 267). At least some of the exits are because lobbyists cultivate familiarity and connections with politicians over time, and these connections lose their value when political power changes hands (Eggers 2010; Vidal et al 2010). Hence, if earmarks are indeed responsive to lobbying (rather than to other factors) we would expect to see a structural break in this relationship as old connections between lobbyists and Congress were disrupted.[21] If on the other hand, our previous estimates were spuriously driven by changes in unobserved disease-specific

[20] We also estimated models separately by chamber, finding the effects of lobbying on earmarks is 1.5 times stronger in explaining Senate earmarking than House earmarking.

[21] Texas Congressman Tom DeLay, reportedly chastised interest groups for sending Democrats to lobby him; "We're just following the old adage of punish your enemies and reward your friends," Delay explained the Washington Post (from Eggers 2010, 11) .

characteristics, then the 2001 change in administration would not have a systematic effect on the returns to lobbying.

Table 4 here

Table 4 presents estimates of the effect of lobbying on earmarks for each of the years during 1999-2008. (We produce these estimates by interacting lobbying with indicators for each year.) The table shows that while the estimated elasticity of earmarks (with respect to lobbying) is in the range of 0.046-0.06 percent for years other than 2002-03, the elasticity drops to 0.037 percent and 0.034 percent for the two years following the 2001 change.

Figure 5 here

Figure 5 plots the estimated elasticities for each year and shows that the elasticity drops from 0.055 in 2001 to 0.037 in 2002 and 0.034 in 2003, but returns to 0.054 in 2004. The estimated elasticities for 2002-03 are significantly different from the elasticities for 2001 and 2004 respectively (at $p<0.05$). We focus on the change in 2001 because both control of the Presidential office and the Senate changed hands during the year, and our study period makes it possible to examine the returns to lobbying both for a reasonable span before and after the change. However, Figure 5 also suggests a drop in the effects of lobbying in 2007 and 2008. Though this change is not as sharp, it coincides with another change in leadership: control of both the House and the Senate (and the chairmanship of the corresponding appropriations committees in charge of NIH allocations) changed from the Republicans to the Democrats in 2007.[22]

These changes in party control and their potential effect on lobbyists are unlikely to be correlated with changes in omitted disease-specific characteristics. While not dispositive of identification questions, they provide evidence at least consistent with a causal interpretation of the estimates of lobbying on earmarking reported in the previous section.

4.3 Favoring powerful groups or responding to information?

Previous criticisms of disease group lobbying for earmarks implicitly have a "rent-seeking" orientation: disease groups are getting earmarks to draw attention (and funding) away from important diseases and those with the highest scientific potential. But some of the empirical literature on lobbying also suggests that interest groups have another potential role: to focus scarce Congressional attention on more salient problems (Evans 1996). Under this "informational" role, interest groups are providing policy makers with the information required to make allocation decisions.

We attempt to disentangle these two effects by investigating whether the effect of lobbying on earmarks responds to changes in disease burden or scientific opportunity. Specifically, we test whether the interaction effects of lobbying expenditures and deaths, and the interaction of lobbying expenditures and publications, influence earmarking. Table 5 shows that the effect of lobbying on

[22] However, while changes in the returns to lobbying appear to lag the 2001 shock by a year, changes in the effects of lobbying are concurrent with the 2007 change in control.

earmarks depends on the magnitude of deaths and publications: each of the interaction terms is positive and significant. However, the effect of lobbying alone is not statistically different from zero.

Table 5 here

One interpretation of this finding is that Congress responds to within -disease increases in lobbying only when accompanied by increases in disease burden and/or scientific opportunity, consistent with the view that lobbying influences Congressional priority setting by providing information. This would be consistent with the Rett Syndrome case study we opened with, where disease group mobilization followed a scientific breakthrough. However, this finding is admittedly tentative, recognizing the various limits of our measures of scientific opportunity and disease burden, discussed earlier.

5 NIH funding for rare diseases

5.1 Relationship between earmarks and NIH allocations

Next, we assess the effects of earmarks, and lobbying-related earmarks, on NIH funding for the 955 diseases. It is possible these lobbying and earmarking activities reflect sound and fury, but do not affect actual agency choices. After all, as we noted above, "soft" earmarks are specified in committee reports that, unlike appropriations bills, are not voted on by the full Congress, and do not have the force of the law. Further, since many of them "urge" and "encourage" funding, rather than use more direct language, it is at least plausible that the NIH need not respond to them.

In examining the effects of earmarking on funding, we focus on new NIH grants, rather than continuations. Although allocations for continuations compose the bulk of NIH's expenditures each year (83 percent of all funding over the 1998-2008 period), these continuations are annual disbursements for multi-year projects funded in previous years, and create inertia in NIH funding patterns. We focus on funding via new grants (accounting for about 17 percent of funding) since these are most plausibly responsive to changes in disease-specific attributes.

Column 1 of Table 6 displays OLS estimates obtained by fitting our data to equation 4, the baseline model. The dependent variable in these regressions is NIH funding for a disease in year t. We find that a 1 percent increase in lagged deaths is associated with a 0.24 percent increase in funding, and that a 1 percent increase in publications is associated with a 0.48 percent increase in funding. (Each of these estimates is significant at the 1 percent level.) These two variables (together with year dummies) explain 21 percent of the variation in NIH funding for diseases in our sample.

Table 6 here

Column 2 shows that the effects of deaths and publications are not statistically different than zero when disease-specific fixed effects are included. Inclusion of fixed effects increases the amount of explained variation to 66 percent. Column 3 introduces earmarks for the diseases to the specification with fixed

effects (as in equation 8). Overall, earmarks appear to have no significant influence on the allocation of NIH funds.

Next, we estimate the effect of lobbying and earmarks on NIH funding with a two stage-least squares model. The first-stage of the model estimates the effect of lobbying on earmarks (as in the previous section). The second-stage uses estimates of the earmarks associated with lobbying (obtained from the first-stage) to estimate the effect of lobbying-driven earmarks on NIH funding. Column 4 of Table 6 reports estimates from this model. We find that a 1 percent increase in earmarks associated with lobbying increases NIH funding for new projects for the corresponding rare disease by 2.2 percent (significant at $p<0.05$). This suggests that, unlike soft earmarks in general, the variation in soft earmarking that is associated with lobbying activity has a strong influence on NIH funding patterns.

5.2 Earmarks and grant mechanisms

In this section, we examine how different grant mechanisms respond to earmarking. This is important since a number of participants in and observers of the NIH allocation process (Varmus 1997, IOM 1998, Drazen and Ingelfinger 2003) suggest that in a context where most of NIH funding is for investigator-initiated research project grants (reviewed by external peer reviewers on the basis of their scientific merit) the channels through which NIH funding could be responsive to extra-scientific considerations, including earmarking, are limited (Sampat 2011).

However, as the IOM (1998) and others suggest, some grant mechanisms—including Requests for Applications (RFAs) and Program Announcements (PAs)—are more focused, and may allow the NIH to steer the direction of research in response to Congressional earmarking. RFAs and PAs are solicitations by the NIH for grant applications that address a defined research topic. PAs are used by the NIH to announce its interest in building or enhancing research in specific areas considered to be of high-priority. RFAs, like PAs, also specify an area of research, with suggested approaches to the research topic described in the announcement.[23]

The IOM notes (about RFAs) "Because of their directedness, such mechanisms tend to be specified by Congress in legislation or report language when Congress concludes that NIH should move more quickly to attack a particular disease or other problem" (IOM 1998). Our own review of report language confirms that RFAs and PAs are commonly requested in these earmarks.[24] A recent report by the Government Accountability Office (GAO, 2000), responding to Senator Harry Reid's (D-NV) queries about how the NIH had been responsive to Congressional report language on Chronic Fatigue Syndrome, notes, "NIH develops extramural research on diseases, including CFS, primarily by creating program announcements for grant applications" (21).

[23] Another difference between a PA and RFA is that the former typically is an ongoing solicitation and accepts multiple applications for the defined area of research. RFAs tend to be one-time solicitations by the NIH.

[24] For example, the FY2006 Committee report emphasizes PAs and RFAs as a means to encourage resource on Charcot-Marie-Tooth disease, an inherited neurological disorder: "The Committee encourages NIH to identify new research opportunities on Charcot-Marie-Tooth that could lead to a relevant program announcement or request for applications."

Figure 6 shows that about 30 percent of NIH funding associated with new grants for rare diseases is funded via RFAs and PAs.

Figure 6 here

Given their apparent connection with earmarking in these qualitative accounts, we examined whether RFAs and PAs are differentially responsive to earmarks and lobbying. Table 7 displays the estimated effects of earmarks (both "uninstrumented" and instrumented with lobbying expenditures) on grants for new projects associated with RFAs and PAs. Both earmarks in general, and those associated with lobbying expenditures (predicted from two-stage models) have a positive association with funding through RFAs and PAs (significant at $p<0.05$). Moreover, the estimated effects of both earmarks and instrumented earmarks are larger in magnitude for RFAs/PAs than for other funds. This confirms that this mechanism is particularly responsive to Congressional influence. As with all grants, earmarks related to lobbying activity have a stronger influence on RFA/PA grants than earmarks in general.

Table 7 here

6 Conclusions

In the U.S., debates about the appropriate roles for politics in the allocation of research funding date back at least to the Bush-Kilgore debates, and debates about the roles of special interest groups in shaping public policy to Madison. This paper was motivated by debates about the effects of lobbying by disease interest groups on NIH's funding choices. Tensions between disease interest groups and the scientific establishment were present even in the early history of the NIH. These tensions intensified in the 1990s, as disease groups intensified their direct lobbying activities, attempting to circumvent the peer review process by obtaining Congressional support for research on particular diseases through report language, or soft earmarks. While the effects of these activities have generated considerable controversy and attention, the effects of advocacy on Congressional earmarking, and the effects of these earmarks on funding patterns, had not been established.

We examined this question in the context of lobbying, earmarking, and funding for nearly one thousand rare diseases over a decade. Not only are rare diseases are interesting in their own right, but they also provide a unique laboratory to examine the effects of lobbying, given the feasibility of reliably linking data on diseases to lobbying, earmarks, and funding, among other variables.

The paper has four main findings. First, we find a positive and statistically significant effect of lobbying on Congressional earmarking (after controlling for disease effects, and variation in deaths and scientific activity). Congressional attention to a disease -- in the form of soft earmarks -- does focus on "squeaky wheels" or variation in lobbying activity, much has the IOM Report and NIH establishment had feared in the 1990s. As in most of the literature on lobbying, there is potential threat to identification: changes in lobbying are not exogenous, and it is possible that unobserved variables are driving both lobbying and earmarking activity. Though we cannot rule these out, a quasi-experiment, reflecting changes in political

power and resulting reduction in benefits of lobbying, provides evidence at least consistent with a casual interpretation of our results.

Second, two-stage models show that the variation in soft earmarking related to lobbying is associated with NIH funding choices. We are agnostic about whether the NIH allocation process, left to its own discretion, would generate the first-best outcome. But for those that believe that peer review process is the route to allocative efficiency, our results suggest some cause for concern.

The result on the effects of "soft" earmarks is new. While most of the literature on lobbying focuses on "hard" earmarks, these have never been prominent at the NIH. Moreover, recent initiatives to restrict "hard" earmarking in the federal budget more generally may well shift attention to soft-earmarking across agencies and issues areas, since the latter are not subject to the same disclosure requirements (Nixon 2008, 2010). It is not obvious ex ante that these earmarks ought to matter for agency choices: after all, they are loosely worded and have no binding force. This is an empirical question. Our results suggest that, at least for NIH funding of rare diseases, earmarks associated with lobbying expenditures do influence the agency's funding choices.

A third, more tentative, finding is that the effects of lobbying on earmarks is most pronounced when accompanied by changes in our measures of disease burden or scientific productivity. A perennial topic of interest in political economy literature is whether lobbying is mainly about rent-seeking, or conveying information. One reading of these results is that disease lobbyists are helping focus scarce Congressional attention on high-salience issues (cf. Jones and Baumgartner). However, this finding is tentative since our measures of burden and scientific opportunity are likely measured with error (as discussed below).

Finally, we find evidence consistent with previous descriptions of the NIH allocation process that certain types of grants (RFAs and PAs) are particularly sensitive to influence by politicians and disease groups, suggesting a mechanism through which non-scientific considerations might influence the direction of publicly funded biomedical research. [25]

As usual, there are caveats. While rare diseases are important and interesting, we chose them mainly for practical reasons. Moreover, as we noted above, the NORD and NIH databases define “rare” diseases expansively. While we can think of no reasons why the findings would not generalize to a broader set of diseases, this is an open question. Second, we do not know for certain if all of the lobbying expenditures

[25] Economists recently have applied theories of “induced innovation” to publicly funded research, but have not specified channels through which considerations beyond science (e.g. disease burden, political activity) could influence the NIH (Cutler, Meara, and Richards 2009; Bhattacharaya and Packalen 2009). Schmookler’s (1965) also suggests that public funding agencies can be responsive to demand-side considerations, though he too does not discuss mechanisms. As Ruttan (1978) pointed out in his discussion of agricultural research, applying induced innovation models to public sector agencies—as opposed to private sector firms—requires identifying explicit mechanisms through which bureaucracies can be responsive. Our results on RFAs/PAs point to one such mechanism. The creation of new NIH Institutes, and cross-Institute allocation decisions, are other channels through which Congress can influence the direction of research. See Sampat (2011) for more discussion.

for disease groups are focused on NIH funding, making it difficult to calculate returns to lobbying. Third, there are potential measurement issues. Scientific opportunity is difficult to operationalize, and deaths are an admittedly blunt measure of disease burden. While in our short panel fixed effects absorb time invariant differences across the diseases, we hope to develop better measures going forward. In addition to these issues, it would be interesting to see if our findings on soft earmarking generalize to other agencies and contexts.

Overall, given the small number of diseases that lobby and get earmarks, and the dominance of investigator-initiated grants (which are less influenced by these political activities) the share of overall allocations influenced by these political activities is likely small, at least for rare diseases in our sample. To assess welfare implications, we also would need to know more about the comparative productivity of research induced by earmarking and that of other grants. These issues are the subjects of ongoing research.

References

Armstrong, Elizabeth; Carpenter, Daniel; Hojnacki, Marie. "Whose Deaths Matter? Mortality, Advocacy, and Attention to Diseases in the Mass Media" *Journal of Health Politics, Policy, and Law* (2006): 729-72.

Carol Ashbury. *Orphan Drugs: Medical Versus Market Value* (Lexington Books, 1985).

Baumgartner and Leech 1998. *Basic Interests: The Importance of Groups in Politics and in Political Science.* Princeton: Princeton University Press.

Bhattacharya, Jay and Mikko Packalen. 2010. "Is Medicine an Ivory Tower? Induced Innovation, Technological Opportunity, and For-Profit Versus Non-Profit Innovation" *NBER Working Paper* 13862.

Cook-Deegan, Robert ; McGeary, Michael "The Jewel in the Federal Crown? History, Politics, and the National Institutes of Health." In *History and Health Policy in the United States: Putting the Past Back In*, edited by Rosemary A. Stevens; Charles E. Rosenberg; Lawton R. Burns, 364: Rutgers University Press, 2006.

David M. Cutler, Ellen Meara, and Seth Richards-Shubik, "Induced Innovation and Social Inequality: Evidence from Infant Medical Care," forthcoming in *Journal of Human Resources*, Volume 47, Number 2 (Spring 2012).

de Figueiredo, J. M. and B. S. Silverman (2006). The returns to lobbying: University lobbying efforts and the allocation of 'earmarked' academic grants. *Journal of Law and Economics* 49, 597{626.

Drazen, J. M., and J. R. Ingelfinger. "Grants, Politics, and the NIH." *N Engl J Med* 349, no. 23 (2003): 2259-61

Dresser, Rebecca. *When Science Offers Salvation: Patient Advocacy and Research Ethics*. New York: Oxford University Press, 2001.

Drew, Elizabeth Brenner. "The Health Syndicate: Washington's Noble Conspirators" *Atlantic Monthly 220* (1967): 75-82

Eastman, Peggy (2002) "How Advocacy Groups Are Increasingly Shaping Research Funding" *Oncology Times*: Volume 24 - Issue 3 - pp 88,91-92

Eggers 2010. The Partisan Revolving Door. Unpublished Draft dated May 2010.

Evans, Diana. *Greasing the Wheels: Using Pork Barrel Projects to Build Majority Coalitions in Congress*. Cambridge University Press 2006.

Garber, Alan and Paul Romer. 1996. "Evaluating the Federal Role in Financing Health Related Research" *PNAS* 1993

Gold, Marte David Stevenson, and Dennis Fryback. 2002. "HALYS and QALYS and DALYS, OH MY: Similarities and Differences in Summary Measures of Population Health" *Annual Review of Public Health* 23: 115-134.

Greenberg, Daniel S."Disease Lobbies." Washington Post October 27 1998, Final ed.: A23.

Grossman, G.M. and Helpman, E. 2001. *Special Interest Politics*. Cambridge and London: MIT Press, 2001. Pp. xv, 364. ISBN 0-262-07230-0.

Hegde, D. 2009. Political Influence Behind the Veil of Peer Review: An Analysis of Public Biomedical Research Funding in the US*. Journal of Law and Economics* vol. 52 p.665-690.

Hegde, D., Mowery, D.C., 2008. Politics and funding in the U.S. Public Biomedical R&D System. *Science*, 12/19/08.

Iaryczower, Matias, Pablo Spiller, and Mariano Tommasi 2002. "Judicial Independence in Unstable Environments, Argentina 1935-1998." *American Journal of Political Science* 46 (4): 699-716 (October).

Institute of Medicine. *Scientific Opportunity and Public Needs: Improving Priority Setting and Public Input at the National Institutes of Health*. (1998).

Jones and Baumgartner (2005) *The Politics of Attention: How the Government Prioritizes Problems*. Chicago: University of Chicago Press.

Kaiser 2009. Kaiser, R. (2009), *So Damn Much Money: The Triumph of Lobbying and the Corrosion of American Government*, Alfred a Knopf Inc.

Lichtenberg, Frank R. "The Allocation of Publicly-Funded Biomedical Research," National Bureau of Economic Research Working Paper No. 6601. Cambridge, MA: NBER, June 1998.

McGeary, M. and Smith, P. (2002). Organizational Structure of the National Institutes of Health Background Paper for the National Academies' Committee on the Organizational Structure of the National Institutes of Health, Washington, D.C.

NIH. (2008). The National Institutes of Health. Office of Extramural Research. Peer review process. Located at http://grants.nih.gov/grants/peer_review_process.htm

Nixon, Ron. "Using 'soft earmarks,' Congress keeps pork projects thriving" *New York Times* April 7, 2008.

Nixon, Ron. "Lawmakers Finance Pet Projects Without Earmarks" *New York Times* December 28, 2010.

Peltzman, S. (1989) The Economic Theory of Regulation after a Decade of Deregulation," *Brookings Papers on Economic Activity* M.N. Baily and C. Winston, Eds., pp. 1-59

Sampat, Bhaven. "The Dismal Science, The Crown Jewel and the Endless Frontier: An Economist's Perspective on the NIH" in D. Foray, ed., *The New Economics of Technology Policy*. Edward Elgar (2009)

Sampat, Bhaven. "Disease Categorization and Priority Setting at the NIH" 2011. Working paper.

Schmookler, J. “Catastrophe and Utilitarianism in the Development of Basic Science” in R. Tybout ed. *Economics of Research and Development* Ohio University Press 1965.

Stigler, G.J. 1971 “The Theory of Economic Regulation,” *Bell Journal of Economics and Management Science* 2(1):3-21.

Strickland, Stephen P. *Politics, Science, and Dread Disease: : A Short History of United States Medical Research Policy*. Harvard University Press, 1972.

Toole, AA. 2007 "Does Public Scientific Research Complement Private Investment in Research and Development in the Pharmaceutical Industry?" *The Journal of Law and Economics* 50, no. 1: 81-104.

U.S. General Accountability Office. 2000. “Chronic Fatigue Syndrome: CDC and NIH Research Activities Are Diverse, But Agency Coordination is Limited.” A Report to Harry Reid, U.S. Senate.

Varmus, Harold. 2009. *The Art and Politics of Science*. Norton.

Vidal I Blanes Draca, M. & Fons-Rosen, C. (2010), Revolving Door Lobbyists. Working paper dates July 2010.

Weisbrod, Burton Allen. *Economics and Medical Research*, AEI Studies 373. Washington: American Enterprise Institute for Public Policy Research, 1983

Zeckhauser, Richard. "Some Thoughts on the Allocation of Resources to Bio-Medical Research." *U.S. Department of Health, Education, and Welfare, Office of the Assistant for Planning and Evaluation, Occasional Paper* (1967).

Tables and Figures

Figure 1a: Excerpt from the Congress Appropriations Conference Committee Meeting Report, FY2003 Appropriations (108th Congress, 1st Session House Report #108-10)

Rett syndrome is a neurological disorder seen almost exclusively in females; it affects approximately one in ten thousand live births per year. The conferees are pleased to learn of the discovery of the MECP2 gene as the main cause of this disorder and encourage the Institutes to expand their research efforts to learn how this gene affects other genes and tissues during the development of the nervous system. The conferees also encourage research to develop animal models of the disorder and to study the daily problems that afflict children with Rett syndrome, including autonomic disorders, as well as research on interventions for improved literacy and communication. Because Rett syndrome is a multi-faceted disorder, the conferees encourage NICHD, NINDS, NIDCD, and NIGMS to work in collaboration to maximize the outcomes from investments made in Rett syndrome research.

Figure 1b: NIH funds for Rett syndrome and the "average" rare disease (2001-2004) (in Millions of FY2010 $)

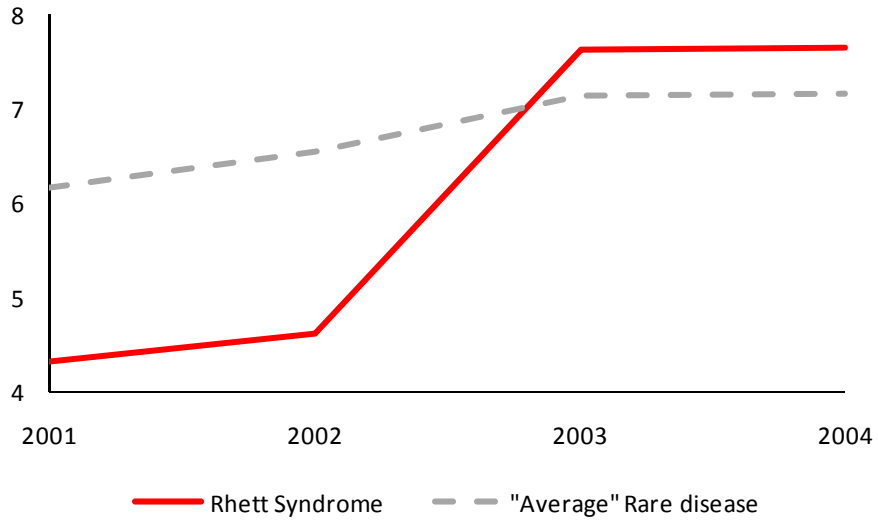


Figure 1b notes: This figure shows NIH funding for research on Rett Syndrome, and the average funding for each of the rare diseases that received NIH grants at least once during the 1998-2008 period.

Figure 2: NIH funds for rare diseases and all other research areas 1998-2008 (in Millions of FY2010$)

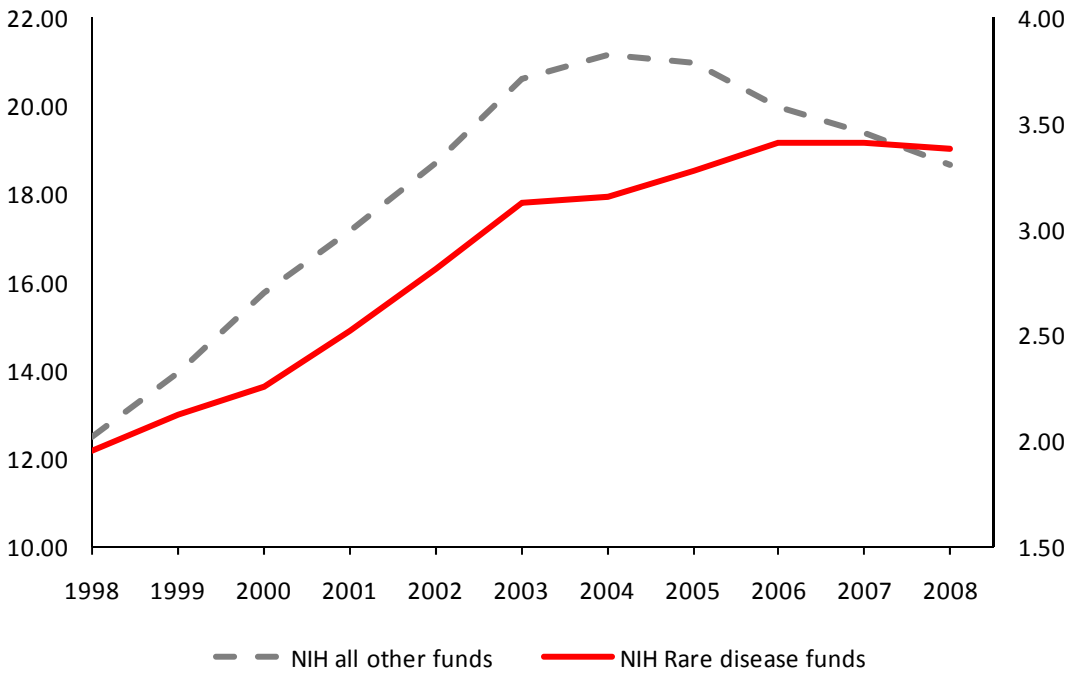


Figure 2 notes: The right axis measures rare disease funding. The left axis measures all other NIH funding.

Figure 3: Rare-disease share of NIH funds, publications and deaths 1998-2007

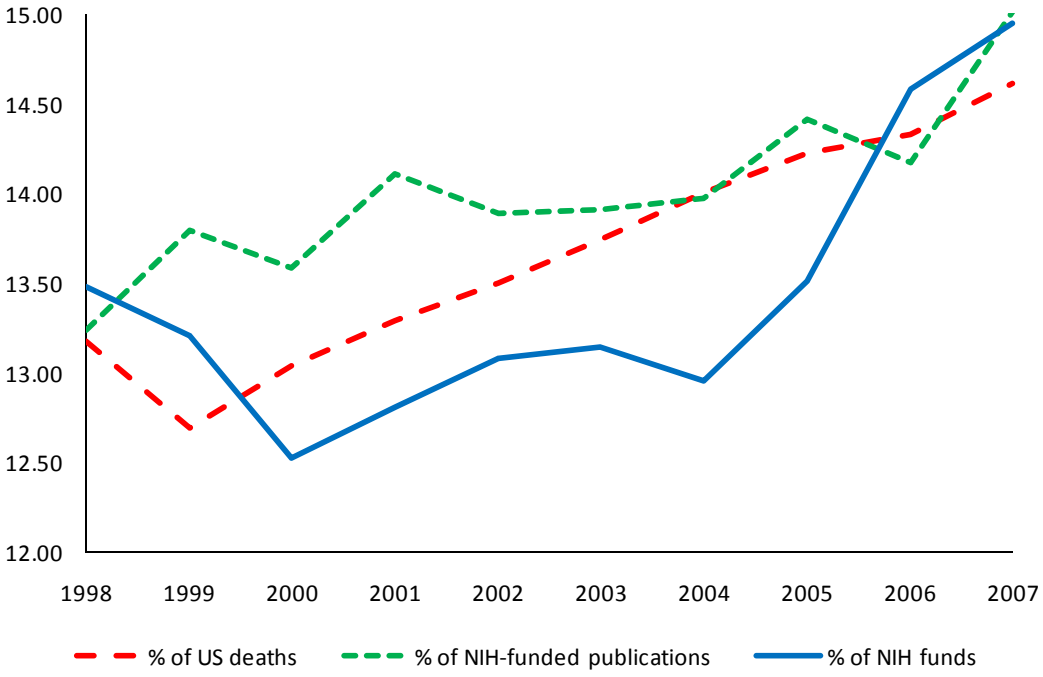


Figure 3 notes: This figure shows the share of US deaths due to the 955 diseases in our dataset, the share of all NIH-funded MEDLINE publications associated with these diseases, and the share of all NIH funds that are for these diseases. The deaths data were collected from the National Vital Statistics System of the National Center for Health Statistics. The publications were extracted from information in raw MEDLINE XML files. (While we used NIH-funded publications only for this chart, it the main text our publications variable measures all publications on a rare disease, regardless of funding source.) NIH funding data were constructed from information in the agency's CRISP and RePORTER databases.

Figure 4: Lobbying expenses of rare-disease interest groups 1998-2008 (in Millions of FY2010$)

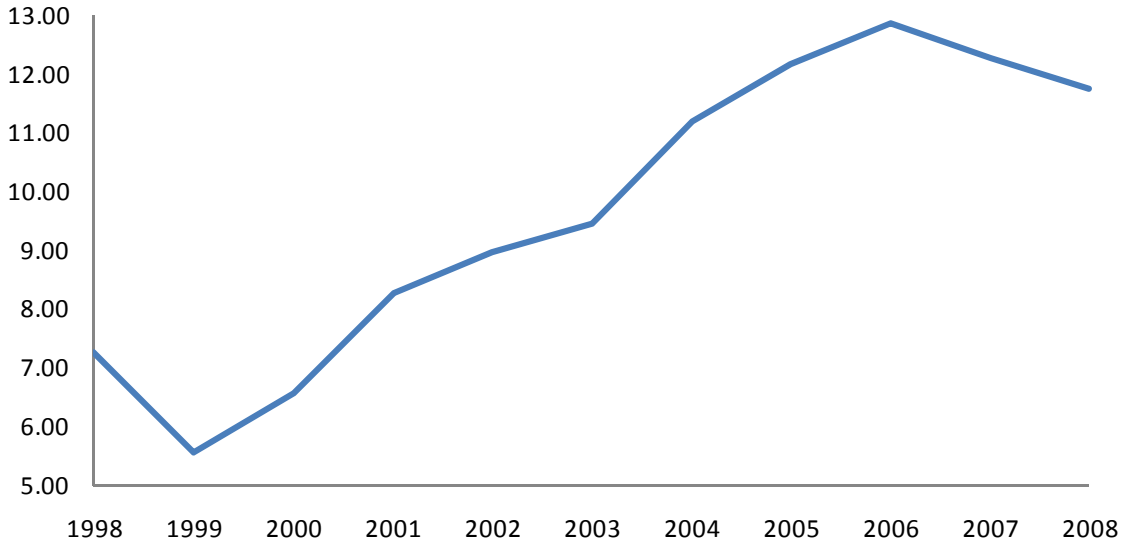
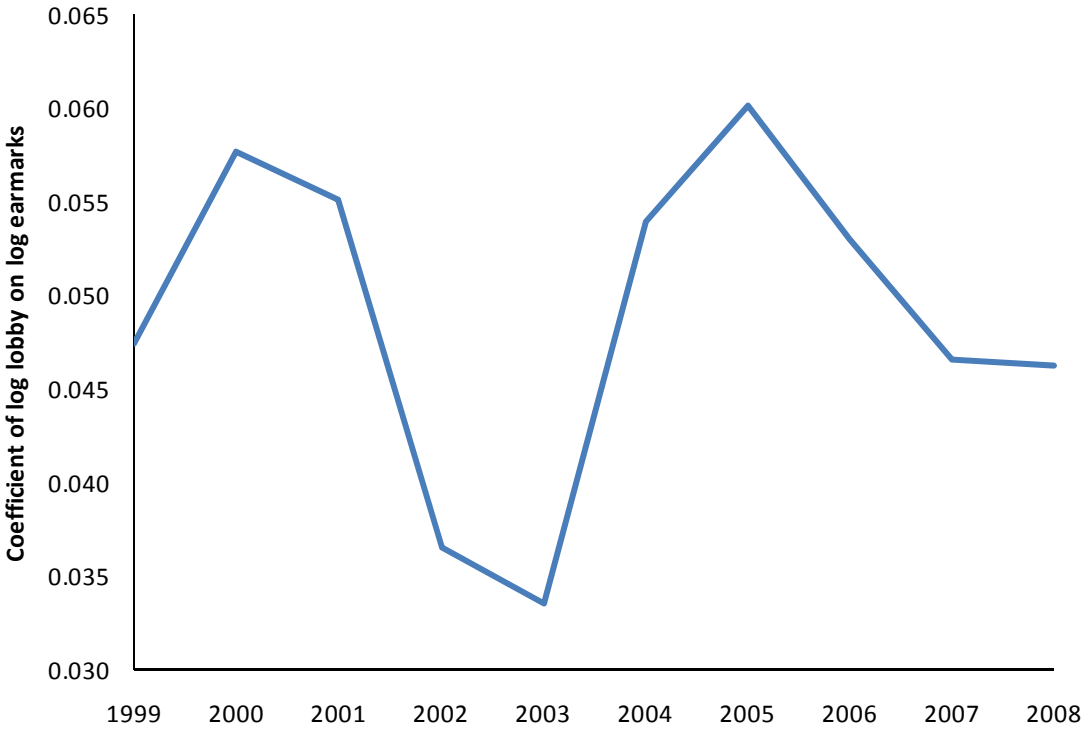


Figure 4 notes: This figure shows the lobbying expenditures of organizations that are identified by the National Organization for Rare Disorders (NORD) as associated with the 955 diseases in our sample. The lobbying expenditures of these organizations were obtained from the Center for Responsive Politics.

Figure 5: Estimated responsiveness of earmarks to lobbying, 1998-2008



Year	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008
President	■	■								
House									■	■
Senate			■	■					■	■

Figure 5 notes: This figure shows estimates of the elasticity of current year earmarking with respect to lagged lobbying, by year. It is based on regression results reported in Table 4. The table below the graph indicates, through shading, years in which Democrats occupied the White House, or controlled the House and/or Senate.

Figure 6: NIH funding (through new grants) for investigator-initiated proposals and RFA/PAs 1998-2008 (in millions of FY2010$)

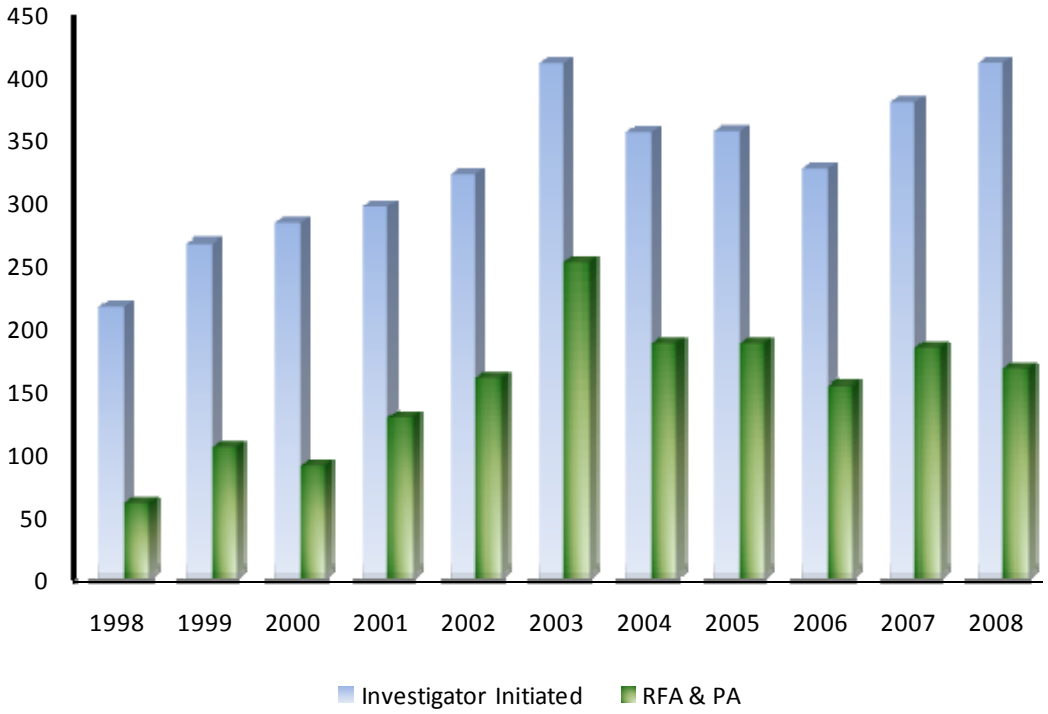


Figure 6 Notes: This figure shows trends in allocations of NIH funds for new applications by the two broad categories of grant-mechanisms at the NIH. Investigator initiated grants are for proposals initiated by researchers. RFAs (Requests for Applications) and PAs (Program Announcements) are solicitations by the NIH for specifically defined areas of research.

Table 1: Deaths, Publications and Lobbying expenditures associated with rare diseases: 1998-2008

Year	Deaths	Publications	Lobbying Expenditures
1998	288,091	15,494	7,232
1999	284,365	16,203	5,548
2000	293,734	16,716	6,536
2001	300,241	18,660	8,244
2002	307,843	19,858	8,958
2003	313,562	20,675	9,437
2004	311,877	22,975	11,189
2005	323,183	23,994	12,168
2006	321,791	25,144	12,843
2007	327,133	18,636	12,267

Table 1 notes: Column 1 shows the number of U.S. deaths associated with the 955 diseases. Column 2 shows the number of publications associated with the diseases. Column 3 shows the lobbying expenditures by disease groups associated with these diseases, in thousands of 2010 dollars. The deaths data were collected from the National Vital Statistics System of the National Center for Health Statistics. The publications were extracted from information in raw MEDLINE XML files. NIH funding data were constructed from information in the agency's CRISP and RePORTER databases. The lobbying expenditures for organizations associated with these diseases were obtained from the Center for Responsive Politics.

Table 2: Descriptive statistics for disease-year observations

Variable	Mean	Std. Dev.	Min	Max
Earmarks (#)	0.09	0.64	0	13.00
NIH grants (in 1000s of FY2010$)	2,991.35	18,572.84	0	488,617.20
Lobbying expenditures (in 1000s of FY2010$)	10.11	37.81	0	1,059.89
Deaths (#)	292.42	2,644.11	0	74,648.00
Publications (#)	18.88	101.11	0	2,770.00

Table 2 notes: This table reports descriptive statistics for the 10,505 disease-year observations in the sample. These are 955 diseases, and we have data on these variables for the eleven year period from 1998-2008.

Table 3: OLS estimates of the relationship between lobbying and earmarks

Dependent variable ("e" represents # of Earmarks)=	log(1+e)	log(1+e)	log(1+e)	e	(e > 0?)
Log Lobby Expenditure (1000's FY2010$; 1-yr lagged)			0.048**	0.207**	0.018**
			[0.009]	[0.039]	[0.006]
Log Deaths (No. of deaths; 1-year lagged)	0.016**	-0.001	-0.002	0.012	-0.004
	[0.002]	[0.005]	[0.005]	[0.014]	[0.004]
Log Publications (No. of publications; 1-year lagged)	0.016**	0.014*	0.012*	0.023+	0.013*
	[0.001]	[0.006]	[0.006]	[0.013]	[0.006]
Log NIH Funds (1000 of FY2010 $; 1-year lagged)			0	0	0
			[0.001]	[0.001]	[0.000]
Constant	-0.017*	0.012	-0.035+	-0.200**	-0.008
Year effects (11 years)	Y	Y	Y	Y	Y
Disease effects (955 diseases)	N	Y	Y	Y	Y
Adjusted-R2	0.065	0.658	0.665	0.605	0.613
N	9,550	9,550	9,550	9,550	9,550
Robust S.E. in brackets; + p<0.1; * p<0.05; ** p<0.01					

Table 3 notes: This table reports Ordinary Least Squares (OLS) estimates of the effect of lagged NIH funding, deaths, and publications on the number and probability of current year earmarks. Column 1 is the baseline model, without disease effects. Column 2 adds disease fixed effects.. Column 3 adds lobbying expenditures of interest groups associated with each disease. The dependent variable in Columns 1-3 is log(1 + earmarks). Column 4 reports estimates of a model with the number of earmarks for a disease-year as the dependent variable. Column 5 reports estimates of a linear probability model with a dummy indicator for whether a disease was subject of an earmark in a disease-year as the dependent variable. We use the full sample of 955 rare diseases and eleven years, but the year 1998 drops out of the estimating sample because our explanatory variables are lagged by a year, and our data begin in 1998.

Table 4: Yearly estimates (OLS) of the relationship between lobbying and earmarks

Dependent variable = Earmarks (e)	log(1+e)
Log Lobby Expenditure X Y1999	0.047**
	[0.011]
Log Lobby Expenditure X Y2000	0.058**
	[0.012]
Log Lobby Expenditure X Y2001	0.055**
	[0.011]
Log Lobby Expenditure X Y2002	0.037**
	[0.010]
Log Lobby Expenditure X Y2003	0.034**
	[0.010]
Log Lobby Expenditure X Y2004	0.054**
	[0.012]
Log Lobby Expenditure X Y2005	0.060**
	[0.011]
Log Lobby Expenditure X Y2006	0.053**
	[0.011]
Log Lobby Expenditure X Y2007	0.047**
	[0.010]
Log Lobby Expenditure X Y2008	0.046**
	[0.010]
Log Deaths (No. of deaths; 1-year lagged)	-0.003
	[0.005]
Log Publications (No. of publications; 1-year lagged)	0.012*
	[0.006]
Log NIH Funds (1000 of FY2010 $; 1-year lagged)	0
	[0.001]
Constant	-0.033
Year effects (11 years)	Y
Disease effects (955 diseases)	Y
Adjusted-R2	0.667
N	9,550
Robust S.E. in brackets; + p<0.1; * p<0.05; ** p<0.01	

Table 4 notes: The table reports ordinary least qquares (OLS) estimates of the effects of lobbying on Congressional earmarking for each of the years during the 1999-2008 period. The coefficients on the interaction terms provide are yearly elasticities.

Table 5: OLS estimates of the relationship between lobbying and earmarks with interaction terms

Dependent variable = Number of Earmarks (e)	log(1+e)	log(1+e)
Log Lobby X Log Deaths		0.006*
		[0.003]
Log Lobby X Log Publications		0.009**
		[0.003]
Log Lobby Expenditure	0.048**	0.002
	[0.009]	[0.006]
Log Deaths (No. of deaths; 1-year lagged)	-0.002	-0.012+
	[0.005]	[0.006]
Log Publications (No. of publications; 1-year lagged)	0.012*	0
	[0.006]	[0.007]
Log NIH Funds (1000 of FY2010 $; 1-year lagged)	0	0.001
	[0.001]	[0.001]
Constant	-0.035+	0.015
Year effects (11 years)	Y	Y
Disease effects (955 diseases)	Y	Y
Adjusted-R2	0.665	0.669
N	9,550	9,550
Robust S.E. in brackets; + p<0.1; * p<0.05; ** p<0.01		

Table 5 notes: This table reports ordinary least squares (OLS) estimates of the effects of 1-year lagged lobbying expenditures on the logged number of current-year earmarks. Column 2 reports estimates of the two interaction terms (a) past-year logged lobbying expenditures and past-year logged deaths, and (b) past-year logged lobbying expenditures and past-year publications, on current year earmarks. Column 1 reproduces the estimates from a specification without the interaction terms, for comparison.

Table 6: OLS and 2SLS estimates of the relationship between earmarks and NIH allocations

DV = Log NIH New Grants (1000 FY2010$)	OLS	OLS	OLS	2SLS
Log (1 + # Earmarks)			0.188	2.211*
			[0.123]	[1.081]
Log Deaths (No. of deaths; 1-year lagged)	0.235**	-0.039	-0.036	-0.038
	[0.016]	[0.077]	[0.077]	[0.074]
Log Publications (No. of publications; 1-year lagged)	0.478**	0.062	0.071	0.064
	[0.016]	[0.069]	[0.069]	[0.074]
Log NIH Funds (1000 of FY2010 $; 1-year lagged)			-0.046**	-0.064**
			[0.014]	[0.015]
Constant	0.264**	1.688**	1.814**	2.422*
Year effects (11 years)	Y	Y	Y	Y
Disease effects (955 diseases)	N	Y	Y	Y
Instrument for Earmarks	N	N	N	Y
Adjusted-R2	0.209	0.692	0.693	0.694
N	9,550	9,550	9,550	8,595
Robust S.E. in brackets; + p<0.1; * p<0.05; ** p<0.01				

Table 6 notes: This table presents ordinary least squares (OLS) and two-stage least squares (2SLS) estimates of the effect of past year (logged) earmarks on (logged) current year NIH funding for diseases. Columns 1-3 provide estimates of the effects of the explanatory variables on all new NIH funding each year. The OLS estimates capture the direct effect of earmarks on NIH grants. Column 4 displays 2SLS estimates of the effects of earmarks on NIH funding. The 2SLS estimation instruments for earmarks with lobbying: the resulting estimate on the earmark coefficient can be interpreted as the effect of lobbying-related earmarks on NIH grants. Since we instrument for earmarks with past-year lobbying, the estimating sample loses an additional 955 observations.

Table 7: OLS and 2SLS estimates of the relationship between earmarks and RFA/PA allocations

D.V. = Log NIH Grants (1000 FY2010$)	OLS	2SLS
Log (1 + # Earmarks)	0.398*	2.801*
	[0.188]	[1.185]
Log Deaths (No. of deaths; 1-year lagged)	0.115*	0.086
	[0.057]	[0.058]
Log Publications (No. of publications; 1-year lagged)	0.092*	0.055
	[0.046]	[0.049]
Log NIH Funds (1000 of FY2010 $; 1-year lagged)	-0.008	-0.013
	[0.009]	[0.009]
Constant	0.277+	0.484
Year effects (11 years)	Y	Y
Disease effects (955 diseases)	Y	Y
Adjusted-R2	0.591	0.586
N	9550	8595
Robust S.E. in brackets; + p<0.1; * p<0.05; ** p<0.01		

Table 7 notes: This table presents ordinary least squares (OLS) and two-stage least squares (2SLS) estimates of the effect of past year (logged) earmarks on (logged) current year NIH funds through RFAs and PAs. The first column displays OLS estimates (which capture the direct effect of earmarks on NIH grants) and the second column reports 2SLS estimates (which instrument earmarks with lobbying and can be interpreted as the effect of lobbying-related earmarks on NIH grants). Since we instrument for earmarks with past-year lobbying, the estimating sample loses an additional 955 observations in the 2SLS regressions.