The Birth of the Congressional Clinic

Raphael Godefroy*

Department of Economics, Stanford University

Abstract

This paper studies the allocation by the US Congress of the National Institutes of Health (NIH) funds for biomedical research across diseases. With both descriptive elements and a model of the allocation process, I argue that the trade-off between productivity concerns and distributive objectives induces Congress to delegate its authority over the allocation decision to the NIH for *basic* research funds, while maintaining control of the allocation of funds across diseases for *clinical* research. I then use the model to estimate the power of the Appropriations Subcommittees on Labor, Health and Human Services of both chambers, which draft the NIH budget. With data on NIH grants and on causes of death in the US for the period 1991-1998, I exploit changes in the composition of these subcommittees to control for research productivity. I find that the NIH budget for clinical research is biased towards diseases that are more critical to the House subcommittee members' constituents. The allocation of basic research funds is not affected by subcommittee composition. In addition, there is no effect of the composition of the Senate subcommittee. Finally, I estimate that the allocation of funds spared 448 deaths due to cardiovascular diseases and 386 deaths due to nervous system and mental health disorders. The counterfactual optimal allocation would have spared 311 and 569 deaths, respectively.

Keywords: Public Health, Government Policy, Publicly-Provided Goods

JEL Classification Number: H4, I1

*I am extremely grateful to Jay Bhattacharya for his generous support and advice. I would also like to thank Tim Bresnahan, Caroline Hoxby, Matt Jackson, Seema Jayachandran, Emily Kopley, Ryan Lampe, Aprajit Mahajan, Neale Mahoney, Eduardo Perez, Julia Tobias and seminar participants at Stanford for helpful comments and guidance. I also thank Dana Goldman, Margaret Blume Kohout, Daniella Meeker who helped me to retrieve part of the NIH data and to become more acquainted with the NIH grants.

1 Introduction

Throughout the last three decades, the public funding of biomedical research has composed a substantial part of the US research budget. In 2008, for instance, the budget of the National Institutes of Health (NIH) was around $28.6 billion, almost 50% of the public budget for non-defense research.[1] In these past few decades, the funding priorities of the NIH have evolved. Between 1991 and 1998, for example, the annual amount of funds for research projects on atherosclerosis has more than tripled, whereas the amount of funds for infectious and bacterial diseases (excluding AIDS/HIV) has increased by 50%. What accounts for these funding shifts? The purpose of this paper is to address this question by studying the allocation of funds across disease-specific research topics. I focus on the period of 1991 to 1998. To be optimal, the amount of funds awarded for research on a disease should take into account the social costs imposed by the disease on the US population, as well as the costs of research production. The contribution of this paper is to show that the allocation of NIH research funds across diseases differs from the optimal solution, controlling for research productivity. In addition, I identify the institutional factor that causes this difference: the bargaining power of congressional subcommittees.

Although the US Congress has authority over the research budget and its allocation across research topics, the NIH is actually in charge of awarding those funds through grants to individual researchers or institutions, after a peer-reviewing process. More importantly, an examination of the formulation of the NIH budget (briefly described below) indicates that Congress delegates at least part of its authority over the allocation decision to the NIH itself. Leaving the management of the research budget to an agency aims to ensure that grants are awarded independently of partisan concerns, as well as to take advantage of the scientific expertise of the NIH's employees. In particular, a large

[1] In comparison, the National Science Foundation budget was $4.5 billion. Source: AAAS *Guide to R&D Funding Data* http://www.aaas.org/spp/rd/guistate.htm

part of those public grants support *basic* research, which encompasses a wide variety of disciplines and methodologies, and should be better evaluated by insiders. NIH-funded programs are not all basic, however. In fact, a substantial part of NIH grants support research projects that apply established methodologies or technologies for the study of human data. Following a usual distinction, these projects are referred to as *clinical* research. The lesser need of expertise to assess the scientific potential of clinical research, in addition to the legal limitations on the use of human data, limit Congress' incentive to delegate the allocation decision to the NIH's experts.

The institutional characteristics of the budget process shed light on a potential source of inefficiency. If Congress directly allocates part of the research funds, these funds will reflect the particular objectives of Congress. What are Congress' objectives? Broadly speaking, they aggregate (1) congressmen's preferences over the allocation of funds across diseases and (2) the distribution of bargaining power among congressmen. For the former, I assume that a congressman's marginal utility from research production on a disease depends on the burden imposed by that disease on her/his constituents. For the latter, I focus on the Subcommittees on Labor, Health and Human Services of the Appropriations Committees of the House of Representatives and the Senate (respectively, SCH and SCS). These subcommittees are in charge of drafting a proposal for the NIH budget to be voted by the whole Congress. I integrate all the factors described in this analysis in a model of the budget process to obtain a formal relation between the allocation of NIH funds and Congress' objectives.

The previous analysis provides two hypotheses for an empirical investigation of the NIH budget. First, clinical research depends more on Congress' aggregated objectives than basic research. Second, the NIH budget favors diseases that are more critical to the subcommittees members' constituents. I test both hypotheses controlling for the productivity of research. To do so, I exploit the changes in subcommittee composition. Following

every congressional elections, seats in these subcommittees are reassigned. Here, I assume that the factors that determine the composition of the subcommittees are independent of the variables of research productivity or congressmen's concerns for medical research. I justify that assumption with evidence from previous studies. With this assumption, I estimate the model for the period of 1991 to 1998. I estimate the disease burdens per district/state and per disease using death records in the Vital Statistics files to estimate congressmen's marginal utility for research. To measure the change in the allocation of funds, I construct a panel dataset using information provided by various NIH databases. This dataset contains, for every year between 1991 and 1998, the amount of NIH funds awarded per research topic and type of research (clinical or basic). It includes as well the number of publications that acknowledge the support of those funds in any of the following seven journals: *Proceedings of the National Academy of Science*, *The Journal of Biological Chemistry*, *The Journal of the American Medical Association*, *Lancet*, *the Canadian Medical Association Journal*, *Nature Medicine* and *the British Medical Journal*.

For the eighteen diseases that have the most substantial impact on the population, I find that the NIH budget for clinical research is biased towards diseases that are more critical to the *House* subcommittee members' constituents. No such effect is observable for basic research. Also, the *Senate* subcommittee appears to have no specific bargaining power. These results suggest that delegation occurs less often for clinical research than for basic research, and that House subcommittee members are more likely to influence the research agenda than other congressmen. An increase in the difference between the burden of a disease borne by the SCH's constituents and by the rest of the country equal to 1% of the national burden increases the amount of funds for clinical research on that disease by 0.64%. In addition, to assess better the consequences of subcommittee power on research production, I use the data on publications to estimate the impact of Congress' objectives on the cost of scientific production, which is measured as the average cost of a publication in the main medical journals. The estimation confirms the greater weight

conferred to the diseases significantly affecting the SCH's members' constituents.

How does subcommittee power impact life expectancy? To address this question, I compare the outcome of the actual funding policy with the counterfactual situation in which the clinical research allocation decision is delegated to the NIH. I estimate the changes in deaths tolls and life expectancy for the whole period per disease, for both policies. To do so, I assume that for a disease of reference, the amount of funds spent for clinical research offsets the gain in number of life years saved. This number is estimated in the literature on the Value of Statistical Life (VSL), which derives from preferences revealed by governmental programs. The estimated outcomes vary substantially across diseases. In particular, I estimate that 180 hundred lives (which represent roughly half as many deaths) lost due to self-destructing behavior or nervous systems disorders would have been saved if clinical research were NIH-controlled. Conversely, committee power on clinical research led to triple the decrease in deaths due to hypertension (24 lives saved for the counterfactual situation; 74 lives saved in the current situation). For all cardiovascular diseases, the optimal allocation would have led to around 120 more deaths.

What follows this introduction is a brief description of the budget process. Section 3 reviews the literature related to this research. The model is detailed in section 4. The data and empirical results are described in sections 6 and 7. The counterfactual analysis is detailed in section 8, and, finally, section 9 concludes.

2 Congress' Authority over the NIH Budget

The congressional documents that refer to Congress' agenda regarding the public policy for medical research suggest that both research productivity and distributive concerns determine the allocation of the NIH budget. In the NIH Revitalization Act of 1996, for instance, the committee for Labor and Health of the senate prescribed the following

policies:

The committee is encouraged by the success of biomedical research in recent years into causes and management of problems associated with diabetes mellitus. However, given the striking increase in the number of patients with diabetes, [...] the committee calls for increased funding in this field so as to encourage new and creative research applications.

Institutionally, Congress decides on the NIH budget every year. The subcommittees on Labor, Health and Human Services of the Appropriations Committees of both chambers are in charge of designing a budget proposal that will then be submitted to the whole Congress - the House of Representatives first, the Senate second. As a whole, Congress controls the amount of public funds devoted to each area of research through specifying the part of the budget to be devoted to every institute composing the NIH or to any specific program operated by the NIH. Some institutes are specifically responsible for supporting research on a single disease or set of diseases (such as the *National Cancer Institute*, or the *National Institute of Allergy and Infectious Diseases*), whereas others may award grants for basic research on any topic (*National Institute of General Medical Sciences*) or support projects studying specific technology (*National Institute of Biomedical Imaging and Bioengineering*). It is to be noted that the existence of these different institutes does not result only from an intrinsic organization of medical research but may be in itself the result of a political bargaining.[2] Furthermore, this budget may include funds targeted to special temporary programs, which allows additional flexibility on the formulation of the budget. Congress does have the institutional power to direct funds towards specific diseases as they have the option to leave the allocation decision for part of the budget to the NIH itself.

[2]For instance, Congress refused the creation of an institute specialized in the funding of research on HIV supported by President Clinton

3 Review of literature

This paper is related to several strands of literature. First of all, it contributes to literature on medical innovations. Much scholarship has studied the determinants of investment in biomedical research. For instance, Cerda[2003] and Acemoglu and Linn [2004] have exploited changes in the market size for a drug caused by the increase in life expectancy to explain innovations in the pharmaceutical industry. Similarly, Kremer[2000] explains that the lack of incentives have led to under-investment in the development of vaccines for diseases that mostly affect poor countries. To my knowledge though, few papers have studied how political concerns may impact the allocation of medical research; the amount of public investment in biomedical research, which is comparable to the amount of industry investment (NIH report[1996]), suggests, however, that political incentives may influence the direction of innovations in the medical sector. One exception is Lichtenberg[2001], who finds a correlation between the allocation of NIH grants and the burden of a disease on the whole US population. His findings are limited, though, by the fact that he considers grants only for the year 1996 and does not control for the productivity of research. Another exception is Bhattacharya and Packalen[2008], who exploit the consequences of aging and obesity on the total population, to show that researchers tend to have biases toward studying more pathologies that have a higher social cost. Like these two last examples, my paper suggests that non-market factors impact biomedical R&D. My main contribution is to identify the institutional characteristics (subcommittee power) that cause this impact. In addition, the analysis over time allows to control for the productivity of research. As such, it completes the investigation of the determinants of innovation that are independent of the "intrinsic interest of research" in Acemoglu and Linn[2004]' words.

This paper also builds upon empirical studies on the effect of political institutions on the distribution of funds. Abundant literature, such as Cohen and Noll[1991], Payne[2003],

tracks pork-barrel funding behind congressional involvement in research and development. However, there seems to be little possibility (and little evidence) of earmarking of NIH grants. All the competing type R grants that I study in this paper, and which represent half of the NIH budget, are awarded after peer-reviewing, which limits the possibility of earmarks. Indeed, for the type of grants I consider, I do not find evidence that the institutions (universities or firms) located in geographic areas represented in the subcommittees receive more funds. My research therefore suggests that committee power can be exerted through other means than the usual pork-barrel phenomenon. In particular, it focuses on the agency problem at the heart of the specific institutional process organizing the design of the research budget.

4 Congress' Delegation Problem

4.1 Model

The purpose of the model is to formalize the Congress' trade-off between productivity and distributive concerns of the Congress. It integrates the institutional facts presented above within a Principal-Agent set-up. The Principal (Congress) has authority over the amount of public funds to be awarded to every area of biomedical research. However, she may decide to delegate the allocation decision for part of the budget to the NIH (Agent), who has more information on the state-of-the-art in biomedical research.

Environment - I define an area of research as the product of a type of research $\theta \in \Theta \equiv \{\theta_B, \theta_C\}$ (corresponding to *basic* and *clinical* research) and a disease $d \in D \equiv \{d1, d2, ..., dn\}$. For a given area of research $\theta d \in \Theta D$, the production of scientific knowledge yielded by an awarded amount $F_{\theta d} \in \Re^+$ is $A_{\theta d} F_{\theta d}^{\gamma}$, $\gamma \in (0, 1)$.

Information - For every $\theta d \in \Theta D$, the productivity factor $A_{\theta d}$ is stochastic, and drawn from a binomial distribution as follows:

$$\forall \theta d \in \Theta D, A_{\theta d} = \begin{cases} 0 & with\ probability\ p_\theta \\ \bar{A}_\theta & with\ probability\ 1 - p_\theta \end{cases}$$

I assume that the NIH observes $A_{\theta d}$, whereas Congress does not.

Preferences - For an allocation plan $(F_{\theta d})_{\theta d \in \Theta \times D} \in \Re^{+\Theta \times D}$, I assume that the expected aggregated utility of Congress C is:

$$U^C((F_{\theta d})_{\theta d \in \Theta \times D}) = E[\sum_{\theta d \in \Theta \times D} A_{\theta d} F_{\theta d}^{\gamma} B_d - m \sum_{\theta d \in \Theta \times D} F_{\theta d}]$$

In addition, I posit that $B_d \equiv \sum_{c \in C} w^c B_d^c$ is the weighted sum of the congressmen' s marginal utilities for scientific production on disease d, where a weight w^c represents the bargaining power congressman c has over the research budget.

The utility of the NIH is:

$$U^{NIH}((F_{\theta d})_{\theta d \in \Theta \times D}) = \sum_{\theta d \in \Theta \times D} A_{\theta d} F_{\theta d}^{\gamma} b_d$$

The marginal utilities B_d^c and b_d are assumed non-negative. In addition, I normalize $\sum_{d=d1...dn} B_d = 1$ and $\sum_{d=d1...dn} b_d = 1$. For both agents, the utility derived from a marginal production of knowledge varies across areas of research, but is similar across types of research on the same disease.

A marginal improvement of the scientific knowledge of a disease would impact every individual affected by the disease. In the rest of the paper, B_d^c will thus be interpreted as a function of the burden imposed by the disease on congressman's c constituents.

Delegation Problem - The general problem of delegation can now be written in the fol-

lowing unfriendly form:

$$\begin{cases} Max_{\tilde{\Theta}\subseteq\Theta,(F_{\theta d})_{\theta d\notin\tilde{\Theta}}\geq 0,F_{\tilde{\Theta}}\geq 0} E[\sum_{\theta d\in\tilde{\Theta}D} A_{\theta d}\tilde{F}_{\theta d}^{\gamma}B_{\theta d}] + \sum_{\theta d\notin\tilde{\Theta}D} A_{\theta d}F_{\theta d}^{\gamma}B_{\theta d} - m\sum_{\theta d\notin\Theta D} F_{\theta d} - mF_{\tilde{\Theta}} \\ such\ that\ (\tilde{F}_{\theta d})_{\theta d\in\tilde{\Theta}D} \equiv \begin{cases} Argmax_{(\tilde{F}_{\theta d})_{\theta d\in\tilde{\Theta}D}\geq 0} U^{NIH}((F_{\theta d})_{\theta d\in\Theta D}) \\ \sum_{\tilde{\Theta}d}\tilde{F}_{\theta d} = F_{\tilde{\Theta}} \end{cases} \end{cases}$$

4.2 Results

The solution of the model provides two results that inform our question. First, in equilibrium, the following necessary condition holds.

Prop. 1: For a given expected productivity, Congress is more willing to delegate the allocation decision for a type of research as the uncertainty on the productivity factor of this type of research increases. With the previous notations, if $\bar{A}_{\theta_C} < \bar{A}_{\theta_B}$ and $(1-p_{\theta_C})\bar{A}_{\theta_C} = (1-p_{\theta_B})\bar{A}_{\theta_B}$, there is no equilibrium such that θ_C is delegated and θ_B is not. The three other possible configurations may arise in equilibrium.

Proof. For the sake of clarity, I consider only two diseases $D \equiv \{d1, d2\}$ in the proof. The general proof for an arbitrary number of diseases is similar. I assume, without loss of generality, that $b_1 > b_2$. Consider first the case in which only one type of research exists. In this case, for a research budget normalized to one, the maximum expected utility attainable for Congress without delegation is equal to (the subscripts for the type of research are dropped):

$$(1-p)\bar{A}B_1\left(\frac{B_1^{\frac{1}{1-\gamma}}}{B_1^{\frac{1}{1-\gamma}} + B_2^{\frac{1}{1-\gamma}}}\right) + (1-p)\bar{A}B_2\left(\frac{B_2^{\frac{1}{1-\gamma}}}{B_1^{\frac{1}{1-\gamma}} + B_2^{\frac{1}{1-\gamma}}}\right)$$

Through delegation, the NIH allocates the whole budget to the most productive topic, and to both topics if both are productive so that Congress's expected utility is then:

$$(1-p)p\bar{A}B_1 + (1-p)p\bar{A}B_2 + (1-p)^2\left[\bar{A}B_1\left(\frac{b_1^{\frac{1}{1-\gamma}}}{b_1^{\frac{1}{1-\gamma}} + b_2^{\frac{1}{1-\gamma}}}\right) + \bar{A}B_2\left(\frac{b_2^{\frac{1}{1-\gamma}}}{b_1^{\frac{1}{1-\gamma}} + b_2^{\frac{1}{1-\gamma}}}\right)\right]$$

The comparison of the two alternatives shows that it becomes more profitable to delegate as p increases, regardless of $\bar{A}$. For two types of research, the separability of Congress' objective function therefore implies that there is no equilibrium such that θ_C is delegated and θ_B is not.

In addition, for a given p and a single type of research,the payoffs obtained in the two cases indicate that it becomes more profitable to delegate as B_1 increases. In particular, there will be no delegation if B_1 is sufficiently close to 0 and delegation if B_1 is in the neighborhood of b_1. For two types of research then, the three other possible configurations (the allocation decision for both types of research is delegated, the allocation decision for basic research only in delegated, no decision is delegated) may arise in equilibrium. □

This proposition states that Congress is more willing to delegate as the informational gap between Congress and the NIH increases. For the type of research is not delegated, the allocation of funds reflects Congress' distributive objectives as stated in the second proposition.

Prop. 2: For the type of research that is not delegated, the amount of funds $(F^*_{\theta d})_{\theta d \in \Theta d}$ has the following form:

$$\forall \theta \notin \tilde{\Theta} \text{ such that } F^*_{\theta d} > 0 : log(\gamma A_{\theta d}) + (\gamma - 1) log(F^*_{\theta d}) + log(\sum_{c \in C} w^c B^c_d)) = m \qquad (1)$$

Proof. The result derives from the first order conditions of the Congress' allocation program. □

This last result can be used to derive a relationship between Congress' objective and funds for research. In what follows, I will assume that the weight of subcommittee members is 'small' compared to one. I thus posit the following distribution of weights across congressmen:

- $w^c \equiv \alpha^H(1 + \frac{435}{435-2\#SCH} w^H)$ if representative c is a member of SCH

- $w^c \equiv \alpha^H(1 - \frac{435}{435-2\#SCH}w^H)$ if representative c is not a member of SCH
- $w^c \equiv \alpha^S(1 + \frac{100}{100-2\#SCS}w^S)$ if senator c is a member of SCS
- $w^c \equiv \alpha^S(1 - \frac{100}{100-2\#SCS}w^S)$ if senator c is not a member of SCS

Here, $\#SCH$ and $\#SCS$ are the number of members in the SCH and SCS respectively. There was no change of congressional rules over the period. It is thus plausible to assume that the sum of the weights, which derives from the institutional characteristics of Congress, has been constant over time. It amounts to assuming that $\alpha^H(435 - w^H) + \alpha^S(100 - w^S)$ is constant. This last formula does not depend on the size of the subcommittees. The coefficients w^S, w^H, α^S and α^H thus can be assumed constant.

The first order expansion of the Congress' objective function then writes, for any disease d:

$$\begin{aligned} log(\sum_{c\in C} w^c B_d^c) &= log[(\alpha^S + \alpha^H) \sum_{c\in House} B_d^c] \\ &+ \frac{\alpha^H}{\alpha^H + \alpha_S} w^H \frac{435}{435 - 2\#SCH} \frac{\sum_{c\in SCH} B_d^c - \sum_{c\in FH} B_d^c}{\sum_{c\in H} B_d^c} \\ &+ \frac{\alpha^S}{\alpha^H + \alpha_S} w^S \frac{100}{100 - 2\#SCS} \frac{\sum_{c\in SCS} B_d^c - \sum_{c\in FS} B_d^c}{\sum_{c\in House} B_d^c} \end{aligned} \tag{2}$$

FH and FS respectively stand for House floor and Senate floor.[3] Note that the sum of burdens imposed by a disease on the congressional districts is equal to the sum of burdens imposed on states, and they are both equal to the national burden of the disease.

Equation (2) displays two variables of interest. I define the *House Relative Burden* of a disease d as:

$$RB_d^H = \frac{435}{435 - 2\#SCH} \frac{\sum_{c\in SCH} B_d^c - \sum_{c\in FH} B_d^c}{\sum_{c\in House} B_d^c} \tag{3}$$

[3]Here the floor refers to all the congressmen who are not in the subcommittee

Similarly, the *Senate Relative Burden* is defined as follows:

$$RB_d^S = \frac{100}{100 - 2\#SCS} \frac{\sum_{c \in SCS} B_d^c - \sum_{c \in FS} B_d^c}{\sum_{c \in Senate} B_d^c} \tag{4}$$

For every disease, the relative burden is proportional to the difference between the share of the national burden of the disease borne by the subcommittee members' constituents and the share borne by the rest of the citizens. This last equation provides a specification to estimate the impact of congressmen' s objectives on the NIH budget, which is the focus of the empirical section of the paper. Here, I have assumed that the total level of budget is endogenous. In particular, an increase of the burden of a disease within a subcommittee may lead to an absolute increase in the amount of funds devoted to that disease. The functional forms presented here would be similar if the budget was exogenous, however, and the specification used in the empirical section can be used regardless of the nature of the budget.

5 Empirical Strategy

The model provides two hypotheses for an empirical investigation of the allocation of the NIH budget. First, a smaller informational gap between Congress and the NIH induces more direct control of Congress over the allocation of funds across diseases. Second, the funds allocated directly by Congress should reflect its objectives, which are a combination of congressmen's preferences and their weights over the allocation decision.

For the first hypothesis, the organization of the NIH allows a distinction between two types of research: basic and clinical. Following the classification of research keywords used by the NIH, clinical research is defined as any research that involves the use of human data. As such, the impact of a clinical research project should be more easily accessible for individuals who have less expertise in the biomedical field than the NIH scientists in charge of awarding research grants. In addition, legal restrictions imposed on

the use of human data (as for instance on the risks of clinical trials) further increase the information a non-expert may have on the potential consequences of a clinical research project, compared to an expert.[4]

For the second hypothesis, the distinction between clinical and basic research may not be enough to identify any impact of Congress on the NIH budget if one does not observe the marginal productivity of research. To control for this, I examine the allocation of funds over a period of eight years, which correspond to four Congress terms. After every congressional election, the composition of the subcommittees changes. These recompositions cause changes in the relative disease burden variables, provided that the burden imposed by a disease varies across districts/states, and provided that Congress has an influence. Much scholarship addresses the question of the endogeneity of congressional committees. For instance, Aghion, Boustan and Hoxby[2005], Knight[2004] and Payne[2003], assume that the composition of the Appropriations committees as a whole is exogenous, and support this assumption with descriptions of the assignment process. The main factors determining committee composition are: seniority, congressmen's influence in their parties, and the concern to have all geographic areas represented. In this paper, I assume that these same factors determine the composition of the subcommittees. It is reasonable to assume that these factors are independent of the productivity of a type of research on a disease, which can then be treated as a fixed effect.

6 Description of the Data

6.1 Data on Disease Burdens

The variables measuring the impact of diseases in the US are constructed using the Vital Statistics files for the years 1991 and 1998. These files contain information from

[4] As an aside, note that if there were a possibility to direct funds towards specific institutions, this might undermine my approach. Obtaining more funds for the institutions that are located in her district/state would represent a simpler way for a congressman to favor her constituents. Although this is not the main question of the paper, I check whether a seat in a subcommittee impacts the amount of research funds allocated to the congressman's state.

individual US death certificates on the state of residence, the age of deceased and the main cause of death. Following the literature on the social costs of diseases, I define the burden of a disease on a state s in a given year as the sum of the deaths of state s's residents during the year, weighted by 100 minus the age of the deceased.

$$B_d^s \equiv \sum_{i \ represented \ by \ s} (100 - Age \ of \ death(i)).1_{i \ died \ of \ d} \tag{5}$$

To focus only on the diseases that represent a substantial burden, I removed from the sample the diseases with a very small effect and partitioned the remaining set of diseases into 18 major causes of death. Altogether, they represent more than 70% of the number of deaths for every year. Table 1a displays how causes of death were aggregated; Table 1b report these causes of deaths, their national burden and the total burden of all causes of death in the US.[5]

Vital Statistics does not mention the electoral district of the deceased to estimate the burden of the disease at the district level. I therefore must make the rough assumption that states are homogeneous, and multiply the disease burden at the state level by the proportion of state residents living in a congressional district. The data on congressional districts were retrieved from Adler(2008).

It is worth mentioning that the burden of a disease on a state, relative to other diseases, defined as $\frac{B_d^S(year \ t)}{\sum_{d \in D} B_d^S(year \ t)}$, has been stable over the period 1991-1998. Figure 1 shows that the correlation between this variable in 1991 and 1998 is roughly equal to one. I thus use only the data for the year 1991.[6]

[5]Some causes of deaths (such as *self destructing behavior*) may not correspond to the common meaning of a disease. Since they represent a large number of deaths and are also the topic of a substantial number of research projects funded by the NIH, I include them in the analysis and refers to all those causes as diseases in the whole paper.

[6]The same results are obtained with the data for 1998.

6.2 Construction of the Relative Burden Variable

To construct the relative disease burden variable, I obtained the list of subcommittee members in both chambers over the period as well as the state in which they were elected. Table 2 shows the size of the subcommittees, as well as the relative burden, for every Congress term. Although the size of the House subcommittee has slightly increased over the period, it represents a small number of House representatives (less than 7%) so that the parameter $\frac{435}{435-2\#SCH}$ is roughly equal to one.

Figures 2 and 3 report the relative burden for every disease and every congress term for each chamber (the average relative burden per disease for the whole period is reported in table 3). The variance of the relative burden for any given year is rather small (around 0.0025), suggesting a relatively small geographic heterogeneity in causes of deaths. However, these figures show no disease-specific pattern or outliers and thus provide a good source of variation.

6.3 Data on Public Funding and Scientific Production

Several public agencies participate in the public support for biomedical research, but, the US Public Health Service (PHS) has been the main provider of public funds for medical research and development in the US. In particular, the NIH, which is part of the PHS, has awarded most of those funds. For instance, in 1990, 85% of the public funds for biomedical research were awarded by the PHS. The NIH itself was responsible for 96% of those funds, i.e. 82% of the overall public budget for biomedical research (Trends in US Funding for Biomedical Research[1995]). The NIH funds research through grants of different types, including research projects differing in size and needs. Among research grants, extramural type R grants are the most popular and represent roughly half of the NIH budget each year (NIH Report[2000]). Competing type R grants typically last between 3 and 10 years, have a $500,000 limit per year and pay on average roughly $160,000 a year. I focus on these grants since they support projects with a well defined focus.

Three sources were used to construct a file of NIH funds and publications per disease and type of research.[7] I use two databases to retrieve information on those grants, awarded between 1991 and 1998. A publicly available database from the NIH website[8] provides the list of every grant awarded for every year between 1991 and 2006.[9] For every year, it contains the amount of funds transferred to any grantee, the location of the grantee, as well as a grant identification number, but no description of the nature of the project funded. I keep only the new grants.[10]. Once an applicant obtains a grant, the NIH transfers every year a certain amount of money during the life of the grant. Since the data provided by the NIH only mention the amount of transfer every year (and not the total amount obtained at the beginning of the grant), I focus on the period 1991-1998, so that almost all the NIH funds awarded through these grants in 1998 have been transferred in 2006.

The *Computer Retrieval of Information on Scientific Projects* (CRISP) database, which also includes all the grants awarded by the NIH, contains a description of the research project funded for every grant and a set of keywords attached to it (but no information on the amount).[11] The keywords mention in particular the disease that is the focus of the research, provided the research is not too broad, and whether the research uses human data. According to the taxonomy of scientific keywords used in the CRISP database, I define all the grants mentioning human data as clinical. I exclude all the grants that contain no keyword that can be linked directly to a given disease.

The third source of data is the *PubMed* database.[12] It records information on a large sample of research publications in health-related fields. In particular, it contains records

[7] A detailed description of the construction of these data is available from the author upon request.
[8] http://report.nih.gov/award/award.cfm
[9] The money was thus received in the fiscal years 1992-1999.
[10] The grants with a serial number starting by 1.
[11] CRISP is available online at http://crisp.cit.nih.gov/
[12] http://www.ncbi.nlm.nih.gov/pubmed/

of all the publications in the main medical and biology journals that were published after 1990. For every publication, it provides the year of publication and, if mentioned in the paper, the grant serial number that supported the authors' research. This serial number is usually only partially given, leading to serious selection problems due to the poor quality of the matching between publications and serial numbers.[13] To have a cleaner dataset, I retrieved only the papers published in the *Proceedings of the National Academy of Science*, *The Journal of Biological Chemistry*, *The Journal of the American Medical Association*, *Lancet*, *Canadian Medical Association Journal*, *Nature Medicine* and *the British Medical Journal*. This choice presents two advantages: these are widely cited journals, and they publish papers from all areas of research, which usually mention the whole grant number.

These data were aggregated so as to obtain one unit of observation per disease, per type of research and per Congress term. Each unit of observation contains the total amount of NIH funds awarded through extramural type R grants to a given type of research on a given disease and the number of publications that acknowledge the support of any grant that would be part of a unit of observation for every year following the first year of the grant. Table 4 displays the descriptive statistics for the whole dataset by year, while table 5 displays the descriptive statistics by disease.

As an aside, let me note that I aggregated the previous data by state and Congress term. This set is used to test whether the presence of a congressman from a given state in a committee causes the institutions of this state to receive more grants.

[13] Overall, it appeared that too many publications were attributed to a serial number.

7 Results

7.1 No Evidence of Earmarking

The panel data of funds per state and year described above was constructed to test whether the institutions located in the electoral units of the subcommittee members are more likely to obtain more NIH funds. In this section, I test this assumption using the NIH extramural grant data mentioned before, but aggregated by state. To do so, I use a difference-in-difference analysis with the following specifications:

$$logF_{s,t} = a_t + c_s + \beta^{House} d_{s \in SCH,t} + \beta^{Senate} d_{s \in SCS,t} + \epsilon_{s,t} \tag{6}$$

and:

$$\frac{F_{s,t}}{\sum_{s=1}^{50} F_{s,t}} = a'_t + c'_s + \beta^{House'} d_{s \in SCH,t} + \beta^{Senate'} d_{s \in SCS,t} + \epsilon'_{s,t} \tag{7}$$

There is one observation per state s and Congress term t. The variable $d_{S \in SCH,t}$ is a dummy equal to one if state s is represented in SCH and $d_{s \in SCS,t}$ is equal to one if state s is represented in SCS. The other variables are term dummies and state fixed effects.

The results are displayed in table 6. The coefficients of interests (the βs in the previous equations) are small and not significant. This supports the assumption that the type of grants I consider are not directly allocated to medical institutions located in states represented in the subcommittees.

7.2 Funds Allocation Across Diseases

I now turn to the estimation of the impact of the relative burden on the allocation of funds. The formalization of the problem leads to the following specification:

$$log(F_{d,\theta,t}) = a_t + b_{d,\theta} + \beta_{S0} RB^S_{d,t} + \beta_{S1} RB^S_{d,t} * dClinic_\theta + \beta_{H0} RB^H_{d,t} + \beta_{H1} RB^H_{d,t} * dClinic_\theta + \epsilon_{d,\theta,t} \tag{8}$$

The variables $RB^H_{d,t}$ and $RB^S_{d,t}$ are, respectively, the House and the Senate Relative disease Burdens, defined by the equations (3) and (4). The unit of observation is a disease d, a type of research (θ, which can be either Applied or Basic) and a Congress term t, from the 102^{nd} (1991-1993) to the 105^{th} (1997-1999) Congressional term. Here, $dClinic_\theta$ is a dummy equal to one if the research is clinical.

I use a fixed effects analysis to estimate this equation. The results are presented in table 7. In addition to the variable of interest, the control variables include a dummy for the type of research, two dummies corresponding to two disease categories (cardiovascular diseases and neoplasms), and interactions between those variables and year dummies. To complete the picture, figures 4 through 7 represent the increase of funds with respect to the Relative Burdens for every term and type of research.

All the results display positive coefficients of interest. First, the Relative Burden variable for the House has a significant positive impact on the allocation of clinical research, whereas no variable of interest significantly affects basic research. In the terms of the model displayed above, clinical research allocation is not delegated, and thus reflects both the committee power in the House and the geographic heterogeneity of causes of deaths between the subcommittee members and the rest of the House. Conversely, Congress' aggregated objective does not impact the allocation of basic research funds, which is more likely decided by the NIH.

Interestingly, the coefficient corresponding to the bargaining power of the Senate subcommittee is very small and not statistically significant. This result is consistent with other findings of the studies of the Appropriations committees.[14] Numerically, the size of the House coefficient is quite large; around 64. This could potentially lead to a substantial effect of committee power on research funds (64 per cent increase in funds) if a disease

[14]For instance, see Knight[2005].

uniquely affected a particular district. However, in practice, the geographic heterogeneity of the impact of the diseases retained here is quite low (a standard deviation is roughly equal to 0.0025), and thus the average effect is less than one percent.

As displayed in table 2, the size of the House subcommittee has increased over the period, as did the NIH budget. Although the increase in size was small, it may cause an endogeneity bias.[15] To check the robustness of the results, I estimate the previous equation for the two congressional terms 1993-1994 and 1995-1996, during which the size of the House subcommittee was constant. This period presents an additional advantage because it corresponds to a change of majority party in Congress. Since committee membership is decided within parties, the shift of majority substantially modified the composition of the subcommittee. Table 8 displays the results of this estimation. The value of the coefficients confirm the role of the House subcommittee's distributive influence over public support to clinical research.[16] Indeed, the specific effect of the relative burden on clinical research for 1993-1996 is roughly double the estimation for the whole period, 1991-1998.

The previous result suggests that the party with majority in Congress may be more determinant than subcommittee power. To address this issue, I include three independent variables that correspond to the Relative Burden between a subset of congressmen and the rest of Congress for the three following subsets: the representatives being members of the House majority party, the representatives being members of the House subcommittee and the House majority party, and the representatives being members of the House subcommittee and the House minority party. The results are displayed in Table 9. Interestingly, all coefficients of the relative burden for clinical research are positive. However, none of them is large or statistically significant, except the coefficient corresponding to

[15] It may stem from an underlying trend in research productivity correlated with that increase.

[16] The small number of observations and the presence of an outlier on the right lead to a negative coefficient for the relative burden.

the group of House subcommittee members who are also in the House majority party. There is thus no evidence that belonging to the majority confers in itself any power on the allocation of NIH funds. However, the results also question the influence that representatives of the minority within the House subcommittee have on the budget.

7.3 Impact on the Cost of Scientific Production

The results presented so far do not give information on the actual effect of committee power on scientific production. Using the data on publications, I define a variable to represent the average cost of scientific production: $c_{d,\theta,t} = \frac{F_{d,\theta}}{Npapers_{d,\theta,t}}$ where $Npapers_{d,\theta,t}$ is the number of papers published in the main journals of medicine acknowledging the use of a grant within 6 years of its awarding.

I then estimate the same equation as above with this new dependent variable:

$$log(c_{d,\theta,t}) = a_t + b_{d,\theta} + \beta_{S0} RB^S_{d,t} + \beta_{S1} RB^S_{d,t} * dClinic_\theta + \beta_{H0} RB^H_{d,t} + \beta_{H1} RB^H_{d,t} * dClinic + \epsilon_{d,\theta,t} \quad (9)$$

The coefficients are displayed in table 10. As can be seen, the significance of the coefficients depends on the controls included. However, the positive sign of all these coefficients in every regression confirm the trade-off between scientific productivity and distributive objectives.

8 Policy Analysis

In this section, I compare the outcomes of the actual funding policy for research with the counterfactual situation in which clinical research allocation is decided by an NIH maximizing the US Welfare (i.e. granting equal weight to every district/state). For this purpose, I focus only on clinical research. I assume that both Congress and the NIH have the perfect information on the productivity of research. In addition, I assume that the

Senate carries no weight and imposes $\gamma = 0.97$.

I posit that a gain of one additional year of life for a disease of reference (infectious diseases excluding AIDS/HIV) and the corresponding average age of death, is equal to \$100000. This corresponds to a typical estimation of the value of a life year saved in the literature on the value of statistical life (The number is derived by Ashenfelter[2006]). To link NIH funds to health outcomes, I assume that the gains in life expectancy and/or reductions in deaths exactly offset the amount of the funds awarded for the research on that set of diseases.

With these assumptions in mind, the productivity factor for infectious diseases ($inf.\ dis.$) in a given year is:

$$A_{inf.\ dis.} = \frac{F_{inf.\ dis.}}{10^5 \times B^{US}_{inf.\ dis.} F^{\gamma-1}_{inf.\ dis.}} \tag{10}$$

where B^{US}_d here is the burden for the whole country. (The subscripts for $Clinic$ and terms are dropped.) Solving the Congress' objective then enables me to derive the productivity factor A_d in every period, since:

$$A_d B^C_d F^{\gamma-1}_d = A_{inf.\ dis.} B^C_{inf.\ dis.} F^{\gamma-1}_{inf.\ dis.} \tag{11}$$

where $B^C_d \equiv B^{US}_d \times [1 + 30(1-\gamma) RB_d]$ and $F_{inf.\ dis.}$ are observed. In addition, knowing A_d allows me to solve the counterfactual optimization problem. I derive for both cases the decrease in burden, assuming that the counterfactual budget for clinical research is the same as the actual. Table 11 displays the results of this estimation. They provide the total amount of funds (assuming the population is constant) for both policies. More interesting are the two alternative estimations of the effect of the decrease in burden. First, I assume that the decrease in burden decreased only the death toll, not the average age of death, for every disease. This estimation is relevant for some sets of diseases,

such as the ones grouped under the generic term "infectious diseases". For instance, a vaccine against hepatitis would eradicate that disease but not affect the other ones. Second, I assume that the decrease in burden did not decrease the number of deaths but only increased the life expectancy of the individuals bound to die of that disease. This approach is more relevant for diseases such as AIDS/HIV.

The table shows a large variation of the impact of committee power across diseases. The most striking figure is the difference between the counterfactual and actual decrease in deaths for the category "self-destructing behavior". NIH-controlled clinical research would have saved 255 lives, while the actual funding for research saved 141 lives. Similarly, research on nervous system disorders (such as brain neoplasms, Parkinson's, etc.) has been underfunded, at the cost of 69 lives. Research on endocrine disorders (which include diabetes) has also also been underfunded. Conversely, research on hypertension has been funded over the decision the NIH would have taken. Funding for research on hypertension corresponds to 74 lives saved, roughly three times what would have happened in the counterfactual situation. For most of the other diseases, the actual and counterfactual impacts were close. These are rough estimates, of course. However, the high returns of investment in scientific research in terms of health outcomes, raised in the literature, suggest that the institutional mechanism leading to the design of the NIH budget can have substantial consequences on health conditions in the US. In spite of the relatively small coefficients estimated in the previous section, the distribution of powers within Congress may have substantial consequences on the disparities of health conditions in the US.

9 Conclusion

Congressional authority over the NIH budget can potentially cause political concerns to determine the allocation of public research funds across research topics. However, this

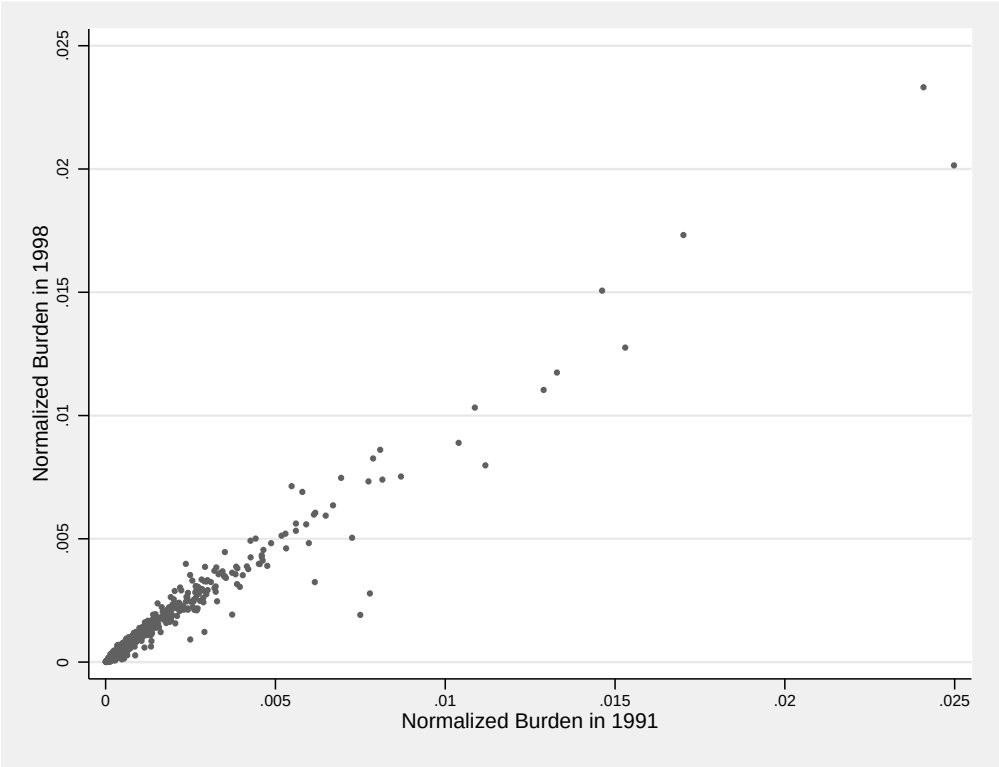
authority may be only formal if the NIH has sufficiently better information regarding the productivity of each topic. The study of the allocation of NIH funds between 1991 and 1998 suggests that the allocation reflects both the geographic distribution of causes of deaths and the distribution of congressmen's influence in the design of the NIH budget in Congress. Focusing on the subcommittees of the Appropriations committees that are in charge of the design of the budget in both chambers, I find that diseases that cause a higher burden on the states or districts that are represented in those subcommittees, relative to the other states or districts, obtain relatively more funds.

These results contribute to the existing literature on the Congress by raising another channel of committee power than the usual pork-barrel. More generally, the issue at stake here is not the pure distribution of private goods among congressmen, already considered in the legislative bargaining literature (as in Baron and Ferejohn[1989] for instance). It is not an ideological issue either, and, as such, requires another approach. The simple formalization presented here relies on mechanisms explored in the Principal-Agent literature. Given that a large part of the public budget is distributed through agencies, these should therefore stimulate a closer investigation, both theoretically and empirically, of the economic effect of these institutions.

Furthermore, the results presented here encourage integration of political factors in the determinants of scientific discoveries in the US. Although the results presented suggest a small effect, on average, of the relative burden, the refinement of the estimation procedure (in particular, including diseases with a smaller impact and having a better measure of the congressmen's actual objectives) should help identify research areas more likely to be affected by Congress' objectives. The exogenous effect of institutions on the public budget for medical research would thus provide an interesting instrument for the study of pharmaceutical and biomedical innovations in the US. This is the goal of further research.

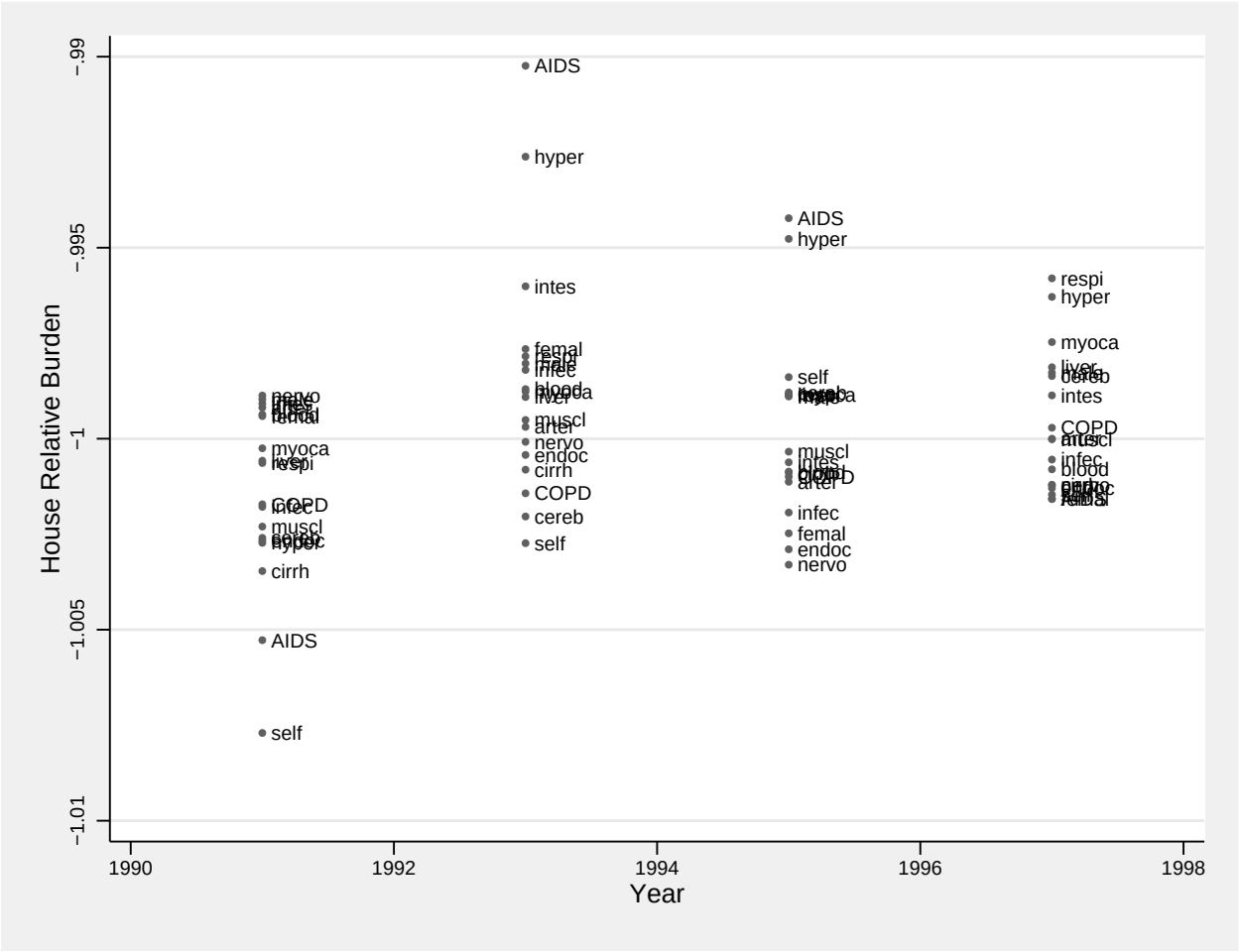
10 Figures and Tables

Figure 1: Comparison of Disease Burdens in 1991 and 1998 per State and Disease



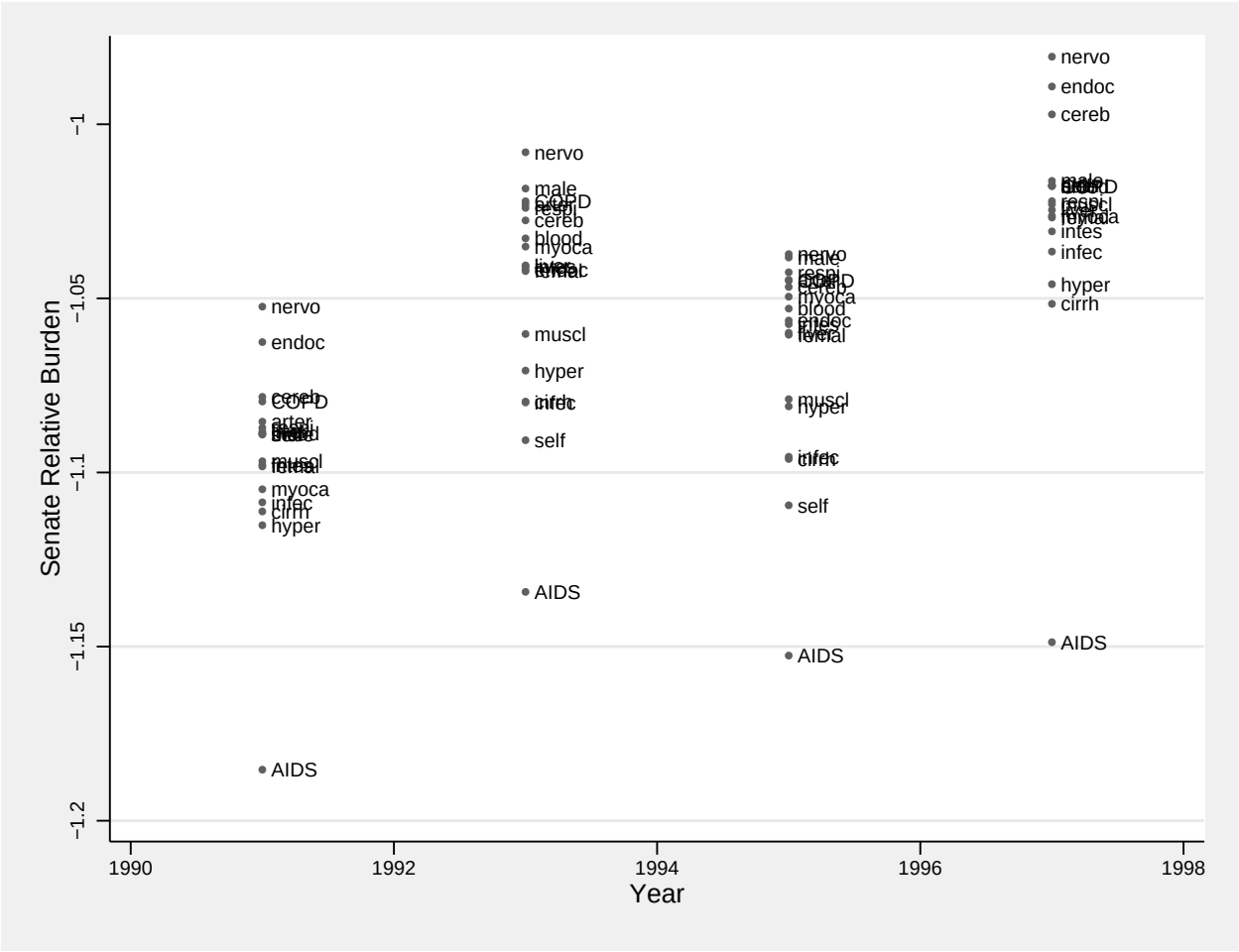
Notes: This figure displays the normalized burdens of a disease and a state in 1991 and 1998, for every disease and US state. For a given state s and disease d, the x-axis (resp. y-axis) reports the normalized burden of d in s in 1991 (resp. in 1998), which is by definition the burden of d in s as defined in equation (5), divided by the sum of burdens over all diseases and states in 1991 (resp. 1998). The figure shows the stability of the distribution of disease burdens across US states over the period 1991-1998.

Figure 2: House Relative Burden per Disease and per Congress Term



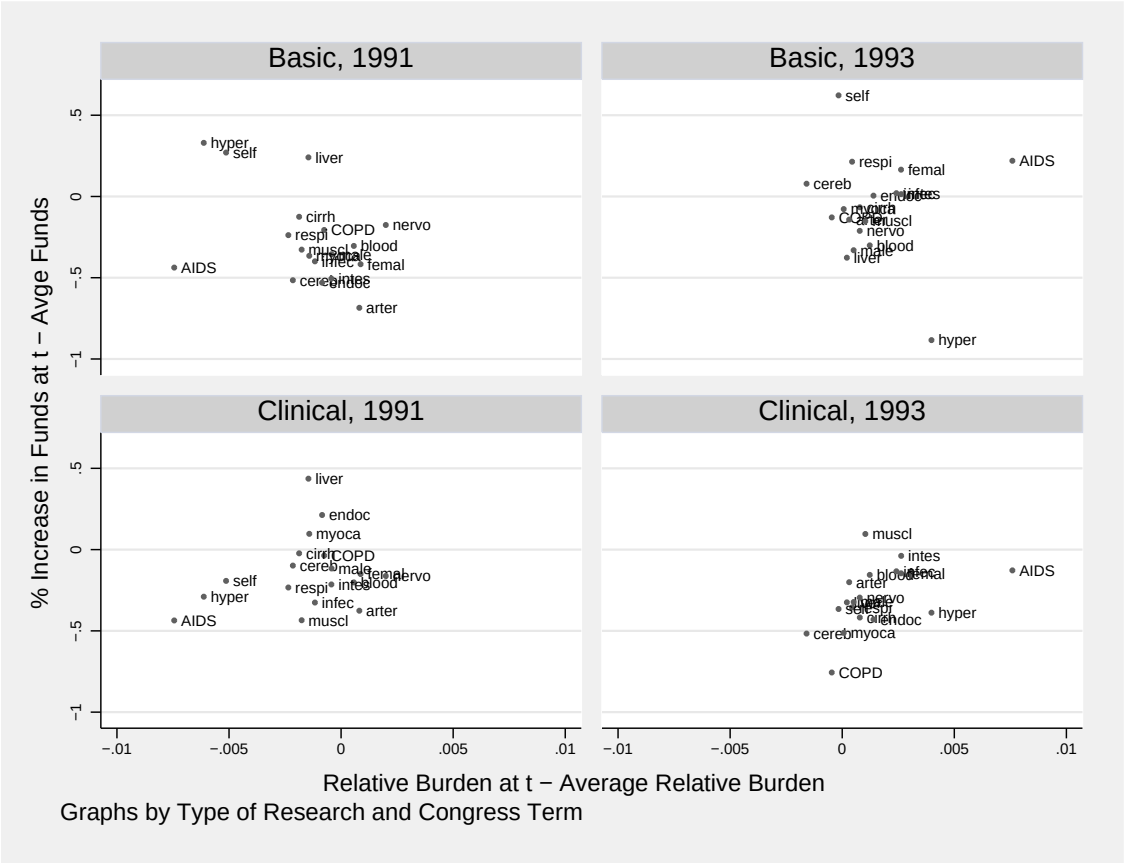
Notes: This figure depicts the *House Relative Burden*, as defined in equation (3) in the text, for every disease and for the four Congress terms: 1991-92, 1993-94, 1995-96, 1997-98. The burden of a disease is defined in equation (5) in the text. The labels indicate the first five letters of the diseases displayed in table 1a.

Figure 3: Senate Relative Burden per Disease and Congress Term



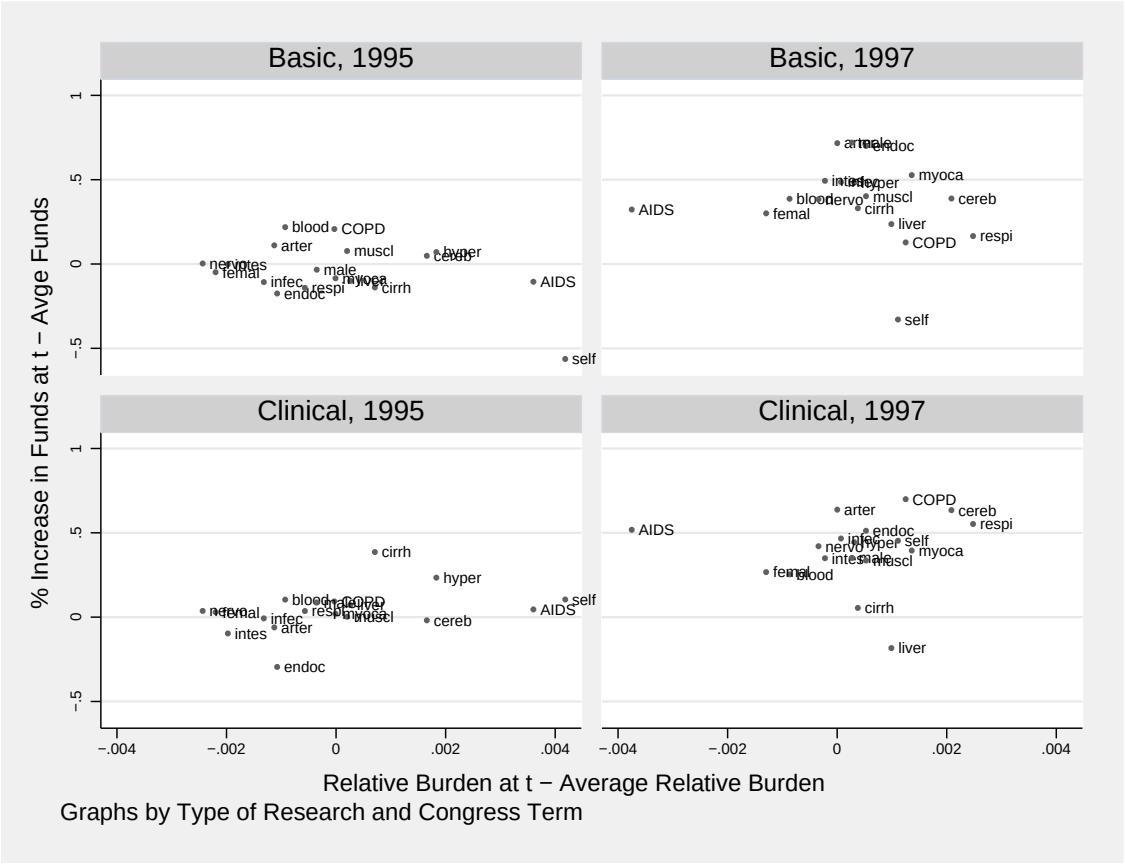
Notes: This figure depicts the *Senate Relative Burden*, as defined in equation (3) in the text, for every disease and for the four Congress terms: 1991-92, 1993-94, 1995-96, 1997-98. The burden of a disease is defined in equation (5) in the text. The labels indicate the first five letters of the diseases displayed in table 1a.

Figure 4: NIH Funds and House Relative Burden
1991 and 1993



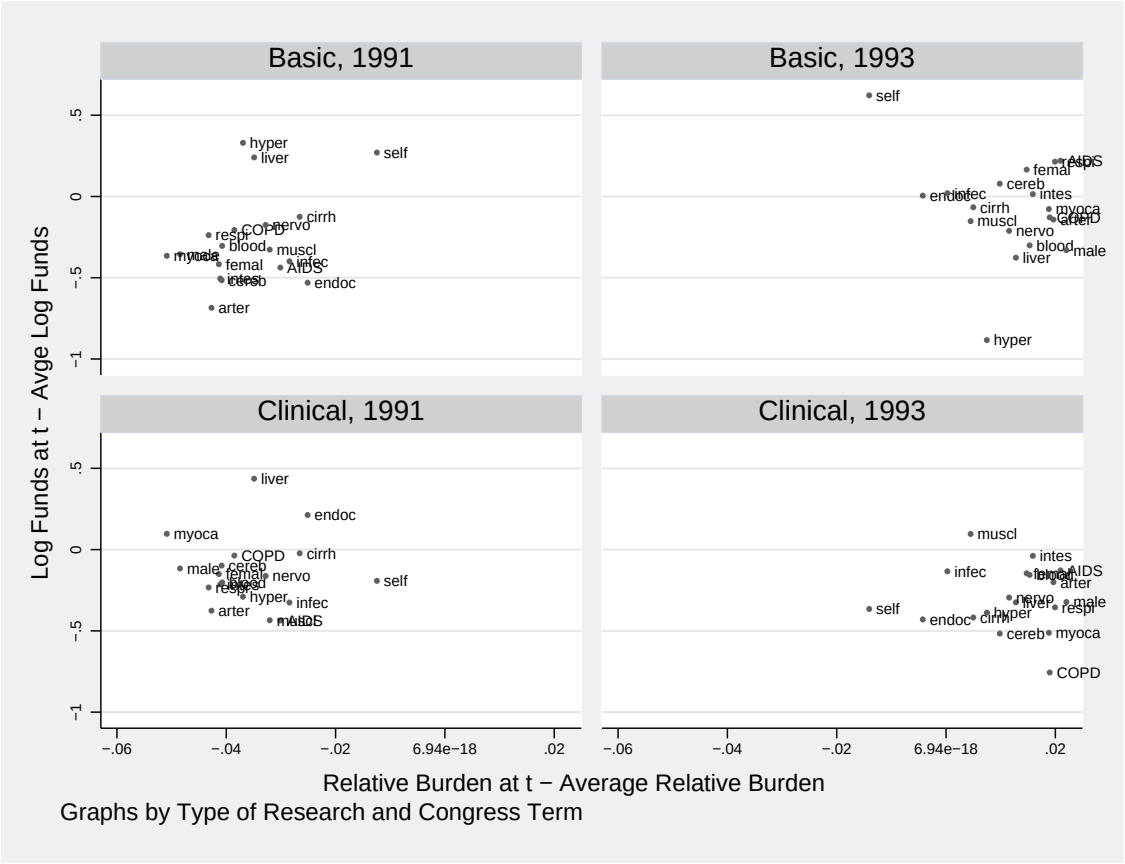
Notes: This figure shows the *NIH funds* (in logs) allocated to a disease with respect to the House Relative Burden of that disease for every disease, as defined in equation (3), for the Congress terms $t = 1991 - 92,\ 1993 - 94$ and for the types of research $\theta = Basic,\ Clinical$. The y-axis reports the difference between the log of funds allocated for research on a disease d during Congress term t, and the average of that variable over the whole period 1991-1998. The x-axis reports the difference between the House Relative Burden of a disease d during Congress Term t, and the average of that variable over the whole period 1991-1998. The *NIH funds* denote the NIH funds awarded through new and competing type R extramural grants.

Figure 5: NIH Funds and House Relative Burden
1995 and 1997



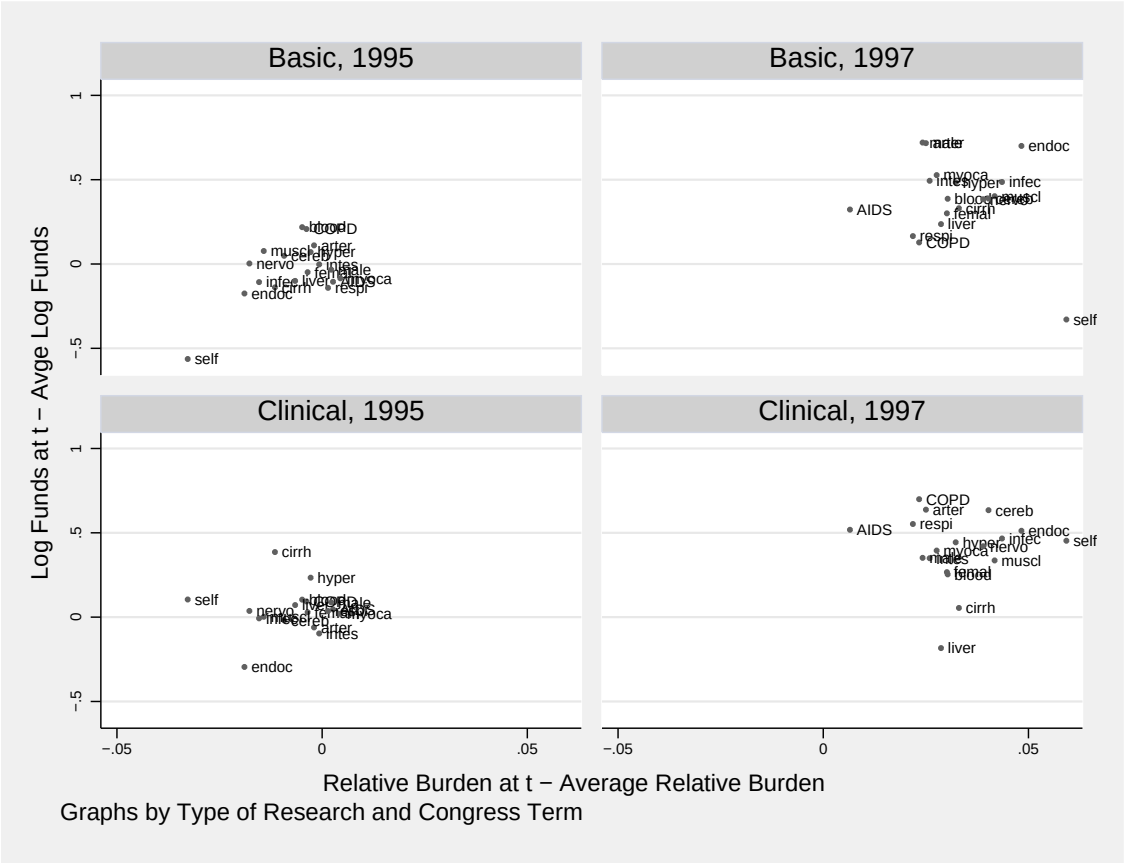
Notes: This figure shows the *NIH funds* (in logs) allocated to a disease with respect to the House Relative Burden of that disease for every disease, as defined in equation (3), for the Congress terms $t = 1995 - 96,\ 1997 - 98$ and for the types of research $\theta = Basic,\ Clinical$. The y-axis reports the difference between the log of funds allocated for research on a disease d during Congress term t, and the average of that variable over the whole period 1991-1998. The x-axis reports the difference between the Senate Relative Burden of a disease d during Congress Term t, and the average of that variable over the whole period 1991-1998. The *NIH funds* denote the NIH funds awarded through new and competing type R extramural grants.

Figure 6: NIH Funds and Senate Relative Burden
1991 and 1993



Notes: This figure shows the *NIH funds* (in logs) allocated to a disease with respect to the Senate Relative Burden of that disease for every disease, as defined in equation (4), for the Congress terms $t = 1991 - 92,\ 1993 - 94$ and for the types of research $\theta = Basic,\ Clinical$. The y-axis reports the difference between the log of funds allocated for research on a disease d during Congress term t, and the average of that variable over the whole period 1991-1998. The x-axis reports the difference between the Senate Relative Burden of a disease d during Congress Term t, and the average of that variable over the whole period 1991-1998. The *NIH funds* denote the NIH funds awarded through new and competing type R extramural grants.

Figure 7: NIH Funds and Senate Relative Burden
1995 and 1997



Notes: This figure shows the *NIH funds* (in logs) allocated to a disease with respect to the Senate Relative Burden of that disease for every disease, as defined in equation (4), for the Congress terms $t = 1995 - 96,\ 1997 - 98$ and for the types of research $\theta = Basic,\ Clinical$. The y-axis reports the difference between the log of funds allocated for research on a disease d during Congress term t, and the average of that variable over the whole period 1991-1998. The x-axis reports the difference between the Senate Relative Burden of a disease d during Congress Term t, and the average of that variable over the whole period 1991-1998. The *NIH funds* denote the NIH funds awarded through new and competing type R extramural grants.

Table 1a: Main Causes of Deaths in the US

Generic Term	*Corresponding Set of Diseases*
AIDS	AIDS, HIV infection
Intestine neoplasm	intestine neoplasm, esophagus neoplasm, stomach neoplasm
Blood/lymphatic neoplasm	blood/lymphatic neoplasm
Respiratory neoplasm	respiratory neoplasm, pharynx neoplasm
COPD	chronic obstructive pulmonary disease, acute bronchitis
Liver neoplasm	liver neoplasm, pancreas neoplasm, peritoneum neoplasm
Cirrhosis	liver cirrhosis, fatty liver, liver circulation disorder
nervous system disorder	nervous system neoplasm, organic brain syndrome, degenerative motor system disease
Muscle/skin neoplasm	muscle neoplasm, skeletal neoplasm, skin neoplasm, skin disorder
Male reprod. disorder	male reproductive system disorder
Female reprod. disorder	female reproductive system disorder, breast neoplasm
infectious diseases (excl. AIDS/HIV)	septicemia, streptococcus infection, enterobacteriaceae disease, rickettsiales disease, actinomycetales infection, staphylococcus infection, influenza, liver infection
Self destructive behavior	self destructive behavior, eating disorder, behavior disorder
Hypertension	hypertension
Endocrine disorder	diabetes mellitus, thyroid disorder
Cerebrovascular disorder	cerebrovascular disorder
Arteries disorders	atherosclerosis, arteriosclerosis, artery occlusion, aortic aneurysm
Myocardial ischemia	myocardial ischemia /hypoxia

Notes: The table reports the 18 (sets of) diseases considered in this paper and the generic term to identify them in the following figures and tables. Two dummies corresponding to two disease categories appear in the following tables:cardiovascular diseases and neoplasms. Cardiovascular diseases include hypertension, myocardial ischemia, cerebrovascular disorders and atherosclerosis.

Table 1b: Main Causes of Deaths in the US - 1991

Disease	#Deaths	US Burden
AIDS	2.8e+04	1.7e+06
Arteries disorders	4.2e+04	9.2e+05
Blood/lymphatic neoplasm	4.5e+04	1.4e+06
Cerebrovascular disorder	1.4e+05	3.1e+06
COPD	9.1e+04	2.3e+06
Cirrhosis	4.4e+04	1.5e+06
Endocrine disorder	5.0e+04	1.4e+06
Female reprod. disorder	6.4e+04	2.1e+06
Hypertension	3.3e+04	8.3e+05
Intestine neoplasm	8.1e+04	2.3e+06
Liver neoplasm	3.9e+04	1.1e+06
Male reprod. disorder	3.5e+04	8.1e+05
Muscle/skin neoplasm	1.4e+04	5.0e+05
Myocardial ischemia	4.8e+05	1.1e+07
Nervous system disorder	4.5e+04	1.1e+06
Respiratory neoplasm	1.6e+05	5.0e+06
Self destructive	5.7e+04	3.4e+06
Infectious diseases	9.9e+04	2.2e+06
Total	1.5e+06	4.3e+07

Notes: This tables reports the number of deaths in the US caused by the diseases reported in Table 1a and the burden of these diseases. The burden is by definition the weighted sum of deaths, the weight being $(100 - age\ of\ death)$, as defined in equation (5).

Table 2: Relative Burden and Size of Subcommittee per Year

Year	Size of SCH	Size of SCS	Relative Burden (+1)	
			House	Senate
1991	12	15	-1.5e-03	-.095
			(2.4e-03)	(.027)
1993	13	15	1.5e-03	-.048
			(3.1e-03)	(.031)
1995	13	15	1.6e-04	-.067
			(2.4e-03)	(.03)
1997	14	15	4.1e-04	-.027
			(1.8e-03)	(.035)
Average	13	15	1.4e-04	-.06
			(2.7e-03)	(.04)

Notes: This table reports the average Relative Burden for both chambers across diseases for every Congress term between 1991 and 1998. The size of a subcommittee is the number of congressmen in the subcommittee. SCH (resp. SCS) stands for the Appropriations Subcommittees on Labor, Health and Human Services of the *House* (resp. *Senate*). Given the small size of the subcommittees, the Relative Burden is close to -1. For the sake of clarity, this table displays 1 + Relative Burden for both chambers. Standard deviations are in parentheses.

Table 3: Relative Burden per Disease

Disease	House Relative Burden (+1)	Senate Relative Burden (+1)
AIDS	2.2e-03	-.155
	(6.3e-03)	(.02)
Arteries disorders	-1.5e-06	-.043
	(7.6e-04)	(.028)
Blood/lymphatic neoplasm	6.8e-05	-.048
	(9.9e-04)	(.028)
Cerebrovascular disorder	-4.4e-04	-.037
	(2.0e-03)	(.031)
COPD	-9.6e-04	-.041
	(8.2e-04)	(.026)
Cirrhosis	-1.6e-03	-.085
	(1.2e-03)	(.024)
Endocrine disorder	-1.8e-03	-.037
	(1.1e-03)	(.031)
Female reprod.	-2.8e-04	-.057
	(2.0e-03)	(.029)
Hypertension	3.4e-03	-.078
	(4.0e-03)	(.027)
Intestine neoplasm	1.4e-03	-.057
	(1.8e-03)	(.027)
Liver neoplasm	8.8e-04	-.053
	(9.6e-04)	(.025)
Male reprod.	1.5e-03	-.04
	(4.3e-04)	(.031)
Muscle/skin neoplasm	-5.4e-04	-.065
	(1.1e-03)	(.029)
Myocardial ischemia	1.2e-03	-.054
	(1.1e-03)	(.033)
Nervous system disorder	-8.6e-04	-.02
	(1.7e-03)	(.029)
Respiratory neoplasm	1.7e-03	-.044
	(1.9e-03)	(.028)
Self destructive behavior	-2.6e-03	-.077
	(3.6e-03)	(.038)
Infectious diseases	-6.2e-04	-.08
	(1.6e-03)	(.029)
Average	1.4e-04	-.06
	(2.7e-03)	(.04)

Notes: This table reports the average Relative Burden for both chambers over time between 1991 and 1998, for every disease. The Relative Burden variables are defined in equations (3) and (4). Given the small size of the subcommittees, the Relative Burden is close to -1. For the sake of clarity, this table displays 1 + Relative Burden for both chambers. Standard deviations are in parentheses.

Table 4: NIH Funds and Publications per Year

Year	NIH Funds			# papers
	All	Basic	Clinical	
1991	1.8e+09	5.1e+08	1.3e+09	3079
1993	1.9e+09	6.8e+08	1.3e+09	3766
1995	2.3e+09	6.7e+08	1.6e+09	3687
1997	3.5e+09	1.1e+09	2.5e+09	4272
Total	9.6e+09	2.9e+09	6.7e+09	1.5e+04

Notes: This table reports the total amount of NIH funds awarded through new and competing type R extramural grants and the number of publications that mention these grants, for every Congress term between 1991 and 1998. The amounts are reported in constant 2000 dollars. # papers is the total number of publications in the main medical journals mentioning the support of grant, per disease and Congress term.

Table 5: NIH Funds and Publications per Disease

Disease	NIH Funds			# papers
	All	Basic	Clinical	
AIDS	1.5e+09	3.4e+08	1.2e+09	1073
COPD	1.4e+08	5.8e+07	7.9e+07	280
Arteries disorder	4.6e+08	1.6e+08	3.0e+08	1222
Blood/lymphatic neoplasm	4.1e+08	1.6e+08	2.5e+08	987
Cerebrovascular disorder	4.6e+08	1.5e+08	3.1e+08	731
Cirrhosis	1.1e+08	5.3e+07	5.8e+07	314
Endocrine disorder	5.7e+08	2.4e+08	3.4e+08	1115
Female reprod. disorder	1.2e+09	3.4e+08	8.9e+08	1796
Hypertension	3.3e+08	1.4e+08	1.9e+08	478
Infectious diseases	7.5e+08	3.6e+08	3.9e+08	1259
Intestine neoplasm	4.0e+08	9.1e+07	3.1e+08	884
Liver neoplasm	1.6e+08	7.2e+07	8.3e+07	464
Male reprod. disorder	3.5e+08	9.7e+07	2.5e+08	588
Muscle neoplasm	2.8e+08	1.2e+08	1.6e+08	744
Myocardial ischemia	3.8e+08	1.5e+08	2.3e+08	490
Nervous system disorder	8.6e+08	2.7e+08	5.9e+08	1654
Respiratory neoplasm	2.8e+08	7.2e+07	2.0e+08	567
Self destructive	9.2e+08	6.5e+07	8.5e+08	158
Total	9.6e+09	2.9e+09	6.7e+09	1.5e+04

Notes: This table reports the total amount of NIH funds awarded through new and competing type R extramural grants and the number of publications that mention these grants, for every disease reported in table 1a. The amounts are reported in constant 2000 dollars. # papers is the total number of publications in the main medical journals mentioning the support of grant, per disease and Congress term.

Table 6: No Evidence of Earmarking

Dependent Variable:	*Log[Funds per state s]*		
	(1)	(2)	(3)
s sits in SCS	-0.026		-0.028
	(0.073)		(0.073)
s sits in SCH		-0.031	-0.033
		(0.065)	(0.065)
Congress Terms Dummies	Yes	Yes	Yes
Constant	0.158	0.158	0.158
	(0.041)**	(0.041)**	(0.041)**
Observations	150	150	150
R-squared	0.08	0.08	0.08

Notes: This table reports the estimation of equations 6 and 7 for the whole period. *s sits in SCH* (resp. *SCS*) means that a representative of state s has a seat in the House (resp. Senate) Subcommittee. There is one observation per state s and per Congress term between 1993 and 1997. Standard errors are in parentheses; * significant at 5%; ** significant at 1%. The coefficients are estimated with a first differences analysis. *NIH funds* denote the NIH funds awarded through new and competing type *R* extramural grants.

Table 7: Allocation of Funds and Relative Burden
1991-1998

Dependent Variable	*Log[NIH Funds per type θ and disease d]*		
	(1)	(2)	(3)
House Relative Burden	-22.207	-35.203	-40.715
	(21.862)	(22.361)	(23.148)
House Relative Burden × Clinical Dummy	37.920	63.912	63.912
	(23.191)	(23.935)**	(23.814)**
Senate Relative Burden	2.074	0.485	0.248
	(3.309)	(4.624)	(4.293)
Senate Relative Burden × Clinical Dummy	-3.534	-0.357	-0.357
	(1.652)*	(4.966)	(4.730)
Congress Term 1993 Dummy	0.011	0.270	0.381
	(0.167)	(0.263)	(0.260)
Congress Term 1995 Dummy	0.199	0.265	0.289
	(0.103)	(0.164)	(0.170)
Congress Term 1997 Dummy	0.577	0.677	0.627
	(0.199)**	(0.318)*	(0.280)*
Clinical Dummy		0.213	0.213
		(0.209)	(0.202)
Congress Term 1993 × Clinical Dummy		-0.519	-0.519
		(0.289)	(0.275)
Congress Term 1995 × Clinical Dummy		-0.132	-0.132
		(0.187)	(0.182)
Congress Term 1997 × Clinical Dummy		-0.201	-0.201
		(0.358)	(0.346)
Cardiovascular Dummy			-0.019
			(0.116)
Neoplasm Dummy			0.020
			(0.103)
Interactions			Yes
Constant	-0.197	-0.303	-0.324
	(0.116)	(0.188)	(0.179)
Observations	144	144	144
R-squared	0.53	0.58	0.62

Notes: The table reports the estimation of equation (8) for the whole period. *House Relative Burden* (resp. *Senate Relative Burden*) stands for the relative burden of the House (resp. Senate) subcommittee with respect to the rest of the House as defined in equation 3. Congress terms are denoted by the year of the first session. *NIH funds per type θ and disease d* denote the NIH funds awarded through new and competing type R extramural grants to research projects of type θ, which can be basic or clinical, studying disease d. The disease categories are cardiovascular diseases and neoplasms. Cardiovascular diseases include hypertension, myocardial ischemia, cerebrovascular disorders and atherosclerosis. *Interactions* means that the interactions of all the dummies of the problem are included in the regression. Robust Standard errors are in parentheses; * significant at 5%; ** significant at 1%.

Table 8: Allocation of Funds and Relative Burden
1993-1995

Dependent Variable:	*Log[NIH Funds per type θ and disease d]*		
	(1)	(2)	(3)
House Relative Burden	-47.900	-51.320	-77.915
	(62.037)	(61.668)	(55.779)
House Relative Burden × Clinical Dummy	102.842	109.681	123.855
	(64.055)	(64.301)	(59.520)*
Senate Relative Burden	13.376	0.584	-30.562
	(17.393)	(31.412)	(20.712)
Senate Relative Burden × Clinical Dummy	-22.951	2.633	27.276
	(8.324)**	(34.801)	(24.734)
Clinical Dummy		0.502	1.208
		(0.696)	(0.562)*
Cardiovascular Dummy			0.723
			(0.351)*
Neoplasm Dummy			0.209
			(0.194)
Clinical Dummy × Cardiovascular Dummy			-0.608
			(0.395)
Clinical × Neoplasm Dummy			-0.339
			(0.229)
Constant	0.227	-0.024	-0.836
	(0.335)	(0.636)	(0.471)
Observations	36	36	36
R-squared	0.26	0.28	0.45

Notes: The table reports the estimation of equation (8) for the 1993-94 and 1995-96 only. *House Relative Burden* (resp. *Senate Relative Burden*) stands for the relative burden of the House (resp. Senate) subcommittee with respect to the rest of the House as defined in equation 3. Congress terms are denoted by the year of the first session. *NIH funds per type θ and disease d* denote the NIH funds awarded through new and competing type R extramural grants to research projects of type θ, which can be basic or clinical, studying disease d. The disease categories are cardiovascular diseases and neoplasms. Cardiovascular diseases include hypertension, myocardial ischemia, cerebrovascular disorders and atherosclerosis. *Interactions* means that the interactions of all the dummies of the problem are included in the regression. Robust Standard errors are in parentheses; * significant at 5%; ** significant at 1%.

Table 9: Allocation of Funds and Relative Burden
Controlling for Affiliation to the Majority Party
1991-1998

Dependent Variable:	*Log[NIH Funds per type θ and disease d]*		
	(1)	(2)	(3)
House Relative Burden for:			
Subcommittee & Majority	-46.724	-51.850	-59.238
	(22.994)*	(24.042)*	(25.227)*
Subcommittee & Majority × Clinical Dummy	64.946	75.129	75.129
	(24.938)*	(26.616)**	(26.372)**
Subcommittee & Minority	35.559	3.744	5.953
	(41.914)	(42.170)	(43.320)
Subcommittee & Minority× Clinical Dummy	-27.457	36.174	36.174
	(47.629)	(48.432)	(48.824)
Majority	-0.407	-0.346	-0.335
	(0.373)	(0.379)	(0.341)
Majority × Clinical Dummy	0.262	0.141	0.141
	(0.464)	(0.485)	(0.502)
Senate Relative Burden for:			
Subcommittee	2.897	0.799	0.681
	(2.856)	(4.235)	(3.837)
Subcommittee× Clinical Dummy	-5.263	-1.068	-1.068
	(1.598)**	(4.704)	(4.522)
Congress Term And Clinical Dummies	Yes	Yes	Yes
Disease Categories Dummies			Yes
Interactions			Yes
Constant	-0.224	-0.331	-0.334
	(0.112)*	(0.181)	(0.172)
Observations	144	144	144
R-squared	0.56	0.59	0.63

Notes: The table reports the estimation of equation (8) for the whole period, where the House Relative Burden for the subcommittee has been replaced by three independent variables corresponding to the relative burden for three sets of representatives: the representatives of the majority party, the representatives of the majority party in the subcommittee and the representatives of the minority party in the subcommittee of the House. The Senate Relative Burden is defined as before. *NIH funds per type θ and disease d* denote the NIH funds awarded through new and competing type R extramural grants to research projects of type θ, which can be basic or clinical, studying disease d. The disease categories are cardiovascular diseases and neoplasms. Cardiovascular diseases include hypertension, myocardial ischemia, cerebrovascular disorders and atherosclerosis. *Interactions* means that the interactions of all the dummies of the problem are included in the regression. Robust Standard errors are in parentheses; * significant at 5%; ** significant at 1%.

Table 10: Cost of Scientific Production and Relative Burden
1991-1998

Dependent Variable	*Log[Average cost of pub. per type θ and disease d]*		
	(1)	(2)	(3)
House Relative Burden	-335.680	-326.135	-332.048
	(200.017)	(186.539)	(190.185)
House Relative Burden × Clinical Dummy	391.632	372.541	372.541
	(211.373)	(187.776)*	(185.649)*
Senate Relative Burden	30.095	51.832	50.226
	(23.202)	(34.850)	(34.692)
Senate Relative Burden × Clinical Dummy	-8.656	-52.130	-52.130
	(10.869)	(35.103)	(34.855)
Congress Term 1993 Dummy	-1.331	-2.355	-1.839
	(0.880)	(1.702)	(1.581)
Congress Term 1995 Dummy	-0.882	-1.956	-2.258
	(0.680)	(1.247)	(1.286)
Congress Term 1997 Dummy	-1.238	-3.110	-3.079
	(1.217)	(2.247)	(2.170)
Clinical Dummy		-1.984	-1.984
		(1.194)	(1.114)
Congress Term × Clinical Dummies		Yes	Yes
Cardiovascular Dummy			0.862
			(0.776)
Neoplasm Dummy			0.202
			(0.776)
Interactions			Yes
Constant	0.863	1.855	1.794
	(0.634)	(1.175)	(1.056)
Observations	144	144	144
R-squared	0.18	0.22	0.30

Notes: The table reports the estimation of equation (9) for the whole period. *House Relative Burden* (resp. *Senate Relative Burden*) stands for the relative burden of the House (resp. Senate) subcommittee with respect to the rest of the House as defined in equation 3. Congress terms are denoted by the year of the first session. *Average cost of pub. per type θ and disease d* denotes the NIH funds awarded through new and competing type R extramural grants to research projects of type θ, which can be basic or clinical, studying disease d, divided by the number of publications mentioning the support of at least one of these grants. The disease categories are cardiovascular diseases and neoplasms. Cardiovascular diseases include hypertension, myocardial ischemia, cerebrovascular disorders and atherosclerosis. *Interactions* means that the interactions of all the dummies of the problem are included in the regression. Robust Standard errors are in parentheses; * significant at 5%; ** significant at 1%.

Table 11: Policy Analysis

Disease	Funds (in millions of $)		Gain in Life Expectancy (in months)		Death Toll Decrease	
	Cfl	Act	Cfl	Act	Cfl	Actl
AIDS	9.9e+02	1.2e+01	4.27	5.03	162	191
Arteries disorders	2.6e+02	3.0e+02	.736	.837	120	136
Blood/lymphatic neoplasm	2.1e+02	2.5e+02	.578	.665	66.6	76.6
Cerebrovascular disorder	2.5e+02	3.1e+02	.214	.256	117	140
COPD	7.3e+01	7.9e+01	.097	.104	28.3	30.4
Cirrhosis	6.8e+01	5.8e+01	.188	.161	19.2	16.5
Endocrine disorder	4.2e+02	3.4e+02	1.02	.816	149	120
Female reprod. disorder	9.6e+02	8.9e+02	1.79	1.68	292	273
Hypertension	6.1e+01	1.9e+02	.222	.68	24.3	74.7
Intestine neoplasm	1.8e+02	3.1e+02	.259	.447	61.6	106
Liver neoplasm	4.6e+01	8.3e+01	.144	.254	16	28.4
Male reprod. disorder	1.3e+02	2.5e+02	.447	.864	55.2	107
Muscle/skin neoplasms	1.4e+02	1.6e+02	1.25	1.38	38.7	42.8
Myocardial ischemia	1.2e+02	2.3e+02	.029	.056	49.7	97
Nervous system disorder	7.6e+02	5.9e+02	2.02	1.57	314	245
Respiratory neoplasm	8.8e+01	2.0e+02	.067	.152	27.8	63.1
Self destructive behavior	1.5e+03	8.5e+02	3.3	1.83	255	141
Infectious diseases	3.9e+02	3.9e+02	.466	.475	175	178

Notes: This table displays the amount of funds in constant 2000 dollars for the actual NIH budget (denoted Act) and the counterfactual (denoted Cfl). Columns 4 and 5 report the corresponding increase in the life expectancy assuming the death toll remains the same. Columns 6 and 7 report the corresponding decrease in the death toll, assuming the average age of death remains the same for those who die of the disease.

11 References

Acemoglu, Daron and Joshua Linn (2004), "Market Size In Innovation: Theory And Evidence From The Pharmaceutical Industry", *Quarterly Journal of Economics,* Vol. 119, No. 3, pp. 1049-1090

Adler, E. Scott (2008), "Congressional District Data File, 1991-1997." University of Colorado, Boulder, CO.

Aghion, Philippe, Leah Boustan, Caroline Hoxby and Jerome Vandenbussche (2005), "Exploiting States Mistakes to Identify the Causal Impact of Higher Education on Growth", NBER Working Paper. Forthcoming, 2005

Aghion, Philippe and Jean Tirole (1997), "Formal and Real Authority in Organizations", *Journal of Political Economy*, Vol 105 No 1.

Ashenfelter, Orley (2006), "Measuring the Value of a Statistical Life: Problems and Prospects", *The Economic Journal*, 116 (March)

Azoulay, Pierre and Abigail Tay (2003), "Medical Progress and Health Care Financing: Research in Academic Medical Centers Following the 1997 Medicare Cuts", Working Paper

Baron, David and John Ferejohn (1989), "Bargaining in Legislatures", *American Political Science Review*, LXXXIII, 1181-1206

Bhattacharyah, Jayanta and Mikko Packalen (2008), "Is Medicine an Ivory Tower? Induced Innovation, Technological Opportunity, and For-Profit vs. Non-Profit Innovation",

NBER Working Paper No. W13862

Cerda, Rodrigo (2003), "Drugs, Market Size and Population" University of Chicago, Ph.D. thesis.

Knight, Brian (2005), "Estimating the Value of Proposal Power", *The American Economic Review,* Vol. 95, No. 5, pp. 1639-1652

Krehbiel, Keith (1990), "Are Congressional Committees Composed of Preference Outliers?" *The American Political Science Review,* Vol. 84, No. 1, pp. 149-163

Lichtenberg, Frank R. (1999), "The Allocation of Public Funds for Biomedical R&D", The AEI Press

Murphy, Kevin M. and Robert Topel (2003), "The economic value of medical research." In: Kevin M. Murphy and Robert H. Topel (eds). *Measuring the Gains from Medical Research: An Economic Approach.* The University of Chicago Press, 2003.

Londregan, John and James M. Snyder, Jr. (1994), "Comparing Committee and Floor Preferences," *Legislative Studies Quarterly*, Vol. 19, No. 2, pp. 233-266

Payne, Abigail (2003), "The Effects of Congressional Appropriation Committee Membership on the Distribution of Federal Research Funding to Universities," *Economic Inquiry*, Vol. 41, No. 2, 325-345