The Birth of the Congressional Clinic*

Raphael Godefroy

Paris School of Economics

July 18, 2011

Abstract

This paper studies the impact of mortality in the districts/states represented in key congressional groups (i.e. committees, subcommittees, and parties) on the National Institutes of Health (NIH) allocation of medical research funds across diseases, for the period 1985-2002. Exploiting the recomposition of any group after congressional elections, I find that congressmen who sit in the House Appropriations Subcommittee on Labor, Health and Human Services, Education and Related Agencies appear successful in steering more funds towards research on diseases that affect their constituents disproportionately. This effect is larger for *clinical* than for *basic* research. No other relevant congressional group, except, to a lesser extent, the House majority, seems to impact that allocation. No group significantly impacts the allocation of funds across states.

Keywords: Legislative Bargaining, Publicly-Provided Goods, Health Policy, Government Policy, Medical Research.

JEL Classification Number: H4, I1

1 Introduction

While a vast literature studies the determinants of private investment in medical R&D (e.g. Acemoglu et al. 2006), little is known on the factors that set the priorities of publicly-supported medical research. The total size of the US federal budget for medical research, however, is comparable to

*I am very grateful to Jay Bhattacharya for his generous support and advice on this project. I also thank Romain Aeberhardt, Laurence Baker, Luc Behaghel, Margaret Blume Kohout, Tim Bresnahan, Joe Doyle, Dana Goldman, Emeric Henry, Caroline Hoxby, Matt Jackson, Seema Jayachandran, Ryan Lampe, Aprajit Mahajan, Neale Mahoney and Daniella Meeker, as well as seminar participants at several institutions for very helpful comments. All errors are mine.

its private counterpart.[1] In addition, abundant normative work emphasizes the value of that public investment (e.g. Arrow 1962, Griliches, Klette and Møen 2000, and Murphy and Topel 2006), and supports its high level.[2] This paper presents a positive investigation of the actual use of that investment, through the study of the allocation of National Institutes of Health (NIH) funds across diseases.

My approach relies on the fact that the NIH agenda partly results from congressional bargaining that takes place throughout the Appropriations process of federal funds. This process involves different congressional groups - committees and subcommittees - that serve well defined institutional roles. I find that congressmen who serve on a single, key, subcommittee, the House Appropriations Subcommittee for Labor, Health and Human Services, Education and Related Agencies (HouS), appear successful in steering NIH funds towards research on disease that affect disproportionately their constituents. This subcommittee, together with the equivalent subcommittee in the Senate, is responsible for designing the proposal for the NIH budget, which will then be discussed by the whole Congress.

Congressional records indeed show that congressmen aim to favor medical research that has higher expected returns in terms of life-years saved. Public funding for research on a disease should therefore reflect the trade-off between the productivity of research on that disease and its impact on mortality. Mortality due to a disease, however, varies across districts and across states, so that it also varies across congressional groups. I use this heterogeneity to estimate the effect of the distribution of life-years lost in the districts/states of any potentially influential group on the distribution of NIH funds across medical conditions. Since the productivity of research on a disease could be correlated with that disease mortality, I use first differenced variables to perform this estimation. The first difference transformation uses the fact that distribution of life-years lost in any congressional group changes every two years, due to the changes in the composition of that group that may follow congressional elections. For that purpose, I use longitudinal data that contain information on mortality for every congressional group and on NIH funds awarded through *type R* grants, at the disease level or at a finer level of aggregation, over the period 1985-2002. Type R grants are the most common NIH grants, and support discrete and circumscribed projects, which can be matched to a well-defined disease category, and represent more than half the NIH budget.

[1] For instance, the National Institutes of Health (NIH) budget and the private investment in pharmaceutical research amounted respectively to around 30 and 40 billion dollars in 2006. See "AAAS Guide to R&D Funding Data," available on http://www.aaas.org/spp/rd/guide.shtml.

[2] As stated by Griliches, Klette and Møen 2000: "There is [...] little controversy among economists about the desirability of governmental support to these activities."

Numerically, I find that an increase of 1 percent of life-years lost because of a disease in the districts represented in HouS increases the funds for research on that disease by 0.8-1.1 percent. On average, this effect corresponds to an increase of yearly funds for clinical research equal to around $2300 for every additional life-year lost in the population of the districts represented in HouS.[3] This effect is much larger for *clinical* research, which, by definition, encompasses any research project that involves human subjects, such as clinical trials, than for *basic* research, which refers to any other project. Life-years lost in no other committee or subcommittee I consider have any robust effect on the allocation of NIH funds.

These results identify mortality in HouS districts as a determinant of the NIH research agenda, as long as diseases affecting more the constituents of HouS congressmen are not systematically more (or systematically less) promising from a medical research perspective. The assignment of congressmen into Appropriations subcommittees is decided within parties, at the beginning of every term, and based on the list of subcommittees, ranked by order of preference, every congressman already assigned to the Appropriations committees submits to his/her party. Although there is no formal congressional rule governing who should sit in some given sub/committee, I couldn't find any evidence suggesting that congressmen were assigned to HouS based on their constituents' health conditions. In fact, HouS members may not have any distinct interest for the NIH budget, or health issues, since the NIH budget is one among several issues they are in charge of, such as labor or education. Instead, congressmen likely rank subcommittees membership according to the potential share of federal funds such membership would let them supervise. The combination of three factors cause this share to vary across subcommittees: the size of the subcommittees, as well as the number of subcommittees a congressman can be assigned to, are bounded by congressional rules, and the total sum of funds any Appropriations subcommittee supervises is mostly exogenous.

If HouS members indeed derive the same utility from medical research as other representatives on average, the previous estimations show that their greater influence on the NIH agenda results from their membership to HouS in itself. These estimations thus provide an insight on the functioning of Congress. This greater influence could stem from the fact that they design the very first proposal that will be discussed in Congress, as in Baron and Ferejohn 1989's theory of legislative bargaining, it may also derive from their interactions with the NIH staff or with medical experts their position

[3]Unless mentioned otherwise, all the monetary amounts in this paper are in dollars of 2000.

involves. If it were not the case, however, those estimations may only reflect heterogeneous preferences for medical research across congressmen. To address this latter concern, I examine the effect of life-years lost in the House majority districts. Due to rules governing the organization of the House Appropriations committees, the share of members of the majority serving on HouS, or on any other Appropriations subcommittee, has to be approximately equal to the share of majority members in the whole House. This institutional feature induces a strong correlation between mortality in HouS and in House majority districts, and, in addition, I estimate that mortality in House majority districts impacts positively the distribution of life-years lost across diseases. This estimation is consistent with the previous results, and is equivalent to using life-years lost in House majority districts as an instrument for life-years lost in HouS. This strong correlation limits the possibility to distinguish the respective of effect of either group, though. To summarize, the affiliation to either HouS or the House majority confers a larger power in the bargaining process on the allocation of NIH funds across diseases.

This paper stresses the limits of the independence of publicly-funded medical research from political influence, which is a major objective of the NIH organization. In fact, congressmen (and the President) have full institutional authority over the federal budget for medical research in the US. Each year, as part of the Appropriations process, they decide on the amount of federal funds allocated to every institute composing the NIH.[4] These institutes are in charge of awarding grants for research on a well-defined set of topics that usually comprises a more or less wide group of pathologies (such as the National Cancer Institute or the National Institute of Dental and Craniofacial Research). Since the institutional rules governing the budget process do not let congressmen to decide on the amount of funds devoted to a disease directly, but only on the allocation of funds across its institutes, congressmen may only affect their *expected* distribution of funds across diseases, conditionally on the distribution of funds across these units. To examine this point, I estimate the effect of life-years lost on expected funds by disease and type of research. The results of this estimation confirm the previous ones. Interestingly, the impact of the House majority is larger in these estimations. Although the House majority has a substantial power to influence the allocation of NIH funds towards institutes,

[4]Two institutions are in charge of funding medical research, development and training specifically: the NIH and the Department of Health and Human Services (HHS) agencies. Both have roughly the same budget, but the HHS supports mainly health services and training, whereas the NIH distributes federal funds to research projects on biomedical topics. Their budget far exceeds the budget of any other R&D agency, except for the Department of Defense agencies. For instance, the budget of the National Science Foundation (NSF), which distributes grants to support research in any scientific field, was about $ 3.6 billion in 2000. Source: The American Association for the Advancement of Science guide to R&D funding data, available on http://www.aaas.org/spp/rd/guihist.htm

HouS thus appears to have more effective means to impact the actual research agenda of the NIH.

Using the same data aggregated at the state level, I then investigate the potential effect of earmarking of NIH funds. I find that the number of representatives from a given state in HouS, or in any other group, has no significant on the sum of funds received by the research institutions located in that state. "Usual pork-barrel", although raised in numerous anecdotes and empirical studies thus seems to have no substantial effect on the allocation of funds across states for the specific type of NIH grants I consider.

The rest of the paper is organized as follows. The next part provides a brief survey of the literature to this paper. The next section inventories the main channels of congressmen's control over the NIH agenda. Section 3 describes the data used for this paper. Section 4 presents the empirical strategy and Section 5 the results of the estimations. Finally, Section 6 discusses the results, focusing on the difference between basic and clinical research, with a model that is the basis of the specification used for the estimations.

Related literature. This paper contributes to two strands of the literature: literature on medical R&D and literature on congressional bargaining. Several studies examine the impact of various sources of funding on the production of medical research (see e.g. Blume-Kohout, Kumar and Sood 2009, Jacob and Lefgren 2007 specifically for the NIH, and the whole field called scientometrics). Productivity of research varies substantially across sources of funding, which is consistent with the fact that motivations other than scientific productivity impact investment decisions. Following the profit incentives hypothesis, Azoulay and Tay 2003, Cerda 2003, Acemoglu and Linn 2004, Lichtenberg and Waldfogel 2009 and Finkelstein 2004, for instance, show that changes in the (potential) demand for a cure, or for a related good, induce changes in the nature or the volume of pharmaceutical innovations. Among them, Azoulay and Tay 2003, Lichtenberg and Waldfogel 2009 and Finkelstein 2004 use public health decisions or recommendations, regarding drugs for orphan diseases and vaccines, as exogenous shocks on the demand of medical goods to control for the productivity of medical research. This paper investigates the very factors driving public decisions themselves. These factors could be used to understand better the complementarity between public and private research (David, Hall and Toole 2000, Hall, Mairesse and Mohnen 2010, Toole 2007), and their joint effect on medical innovations.

Few papers study the factors that impact the amount of public funds devoted to research on some given topic. Lichtenberg 2001 investigates the correlation between NIH funds and mortality. Bhattacharya and Packalen 2008 and Cutler, Meara and Richards 2009 link the level of investment in research to mortality in the US, the former using the same instruments as Acemoglu and Linn 2004. Like Ellison and Wolfram 2001 and Miller 2008, this paper stresses the role of political economy on medical issues. The contribution of this paper to this literature is to show that the link between mortality and public funding of medical research results from a combination of congressional power and local concerns. In addition, this paper contributes to the open question of empirical effects of the composition of majorities in Congress (see e.g. Jayachandran 2006 and Snowberg, Wolfers and Zitzewitz 2006).

The local concerns studied here also differ from the local pork-barrel objectives usually raised in the political economy literature on congressional bargaining. Several studies (e.g. Cohen and Noll 1991, Ferejohn 1974, Atlas et al. 1995 and Knight 2005) show how the geographic origin of members of relevant committees affects the geographic allocation of federal funds. This influence has proven useful in explaining other differences across states (e.g. Aghion et al. 2005 or Levitt and Snyder 1997). For the NIH, Payne 1999, Payne 2003 and Hegde 2009 find some impact of congressional power on the geographic allocation of NIH funds. Indeed, Hegde 2009 finds that research centers located in the states represented in HouS do receive more research funds, but mitigates the impact of such earmarking by showing that this effect is limited to centers with relatively little expertise. The bulk of NIH funds traditionally goes to the same small set of universities, which do have expertise in any field, so that, on aggregate, earmarking to local research institutes should have limited consequences. The results here are consistent with that hypothesis, since I find that no substantial pork-barrel at the state level.

2 Congressional control over the NIH budget

In his memoir *The Art and Politics of Science*, Harold Varmus, former head of the NIH, mentions that he once received a call from the congresswoman Nancy Pelosi,

> "asking [him] to add $50 million for the budget for AIDS research. As the representative from one of the districts most heavily affected by the epidemic, her wishes were understand-

> able. Since she was a member of the House Appropriations Subcommittee for the NIH, she was in a position to try to increase funds for AIDS research when the subcommittee was debating the size of the NIH budget."

He later adds that he "declined as politely as [he] could. Sometimes it was not so easy to say no." (Varmus 2009 p.163.) This example illustrates that congressmen are aware of the diseases affecting their constituents, and that they are active in trying to influence the NIH budget priorities. Although Varmus 2009 indicates that that phone call had little effect on the budget, Hegde 2009 provides several examples suggesting that other similar attempts were successful.

Regardless of the actual success of such attempts, there are direct ways through which congressmen can control, at least partly, the funding for research on a disease.

2.1 Control through the budget formulation

The main channel of congressmen's control on the medical research agenda lies in the formulation of the budget. Every year, the Appropriations committees of both chambers are in charge of designing a budget proposal, based on a budget request from the NIH, that will then be submitted to the whole Congress — the House of Representatives first, the Senate second.[5] The Appropriations bill for the NIH specifies the budget to be received by every institute composing the NIH. These institutes are relatively autonomous units that fund projects on some specific area of scientific research. 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 (such as the National Institute of General Medical Sciences) or support projects studying specific technology (such as the National Institute of Biomedical Imaging and Bioengineering). As a result, in several instances, the area of specialization of the institutes and centers overlap, whereas some topics are strictly included in the area of a single institute.

In addition to setting the budget of each institute, the Appropriations bill may also make additional recommendations on how to allocate funds across research topics. An examination of the bills suggests that both research productivity and distributive concerns determine the allocation of the NIH

[5] Like the rest of the budget, it is enacted unless it is vetoed by the President.

budget. The congressional report of the 104th Congress relative to the Departments of Labor, Health and Human Services, and Education, and Related Agencies Appropriations bill contains the following statement:[6]

> "The [Appropriations] Committee believes that NIH should distribute funding on the basis of scientific opportunity.[...] To enhance NIH's flexibility to allocate funding based on scientific opportunity, the Committee has attempted to minimize the amount of direction provided in the report accompanying the bill. [...] In stating that scientific opportunity should be the basis for allocating research funding, the Committee understands that other factors are also relevant to NIH's decisions, including such considerations as the infectious nature of a disease, the number of cases and deaths associated with a particular disease, the Federal and other costs of treating a disease, the years of productive life lost due to a particular disease, and the estimated proximity to research breakthroughs. The Committee does not presume to judge which criteria should take precedence in individual funding decisions, but urges NIH to consider the full array of relevant criteria as it constructs its research portfolio."

The bills often specify how the funds might be allocated across topics within the institutes or centers. As in the previous excerpt, the bills often repeat that the NIH has authority on the allocation of funds, yet give precise suggestions for the use of those funds. For instance, the previous report mentions:

> "AIDS Funding. — Consistent with the philosophy outlined above, the Committee has again chosen not to earmark a specific dollar amount for AIDS research and has not provided a single appropriation for the Office of AIDS Research. In relying on NIH's recommendations for the allocation of the total funding provided by the Committee, the Committee understands that it would be NIH's intent to allocate AIDS funding in the following manner:[...]" [The report then details the amount of funds for research on HIV to be awarded by every institute or center.]

[6] 104th Congress, 2nd Session House of Representatives, 104-659, Departments of Labor, Health and Human Services, and Education, and Related Agencies Appropriation Report, 1997.

The Appropriations committees of either chamber, cited here, are not the only groups that may influence the appropriations of funds across research areas: within the Appropriations committees, the Subcommittees on Labor, Health and Human Services, Education and Related Agencies of the Appropriations Committee have the specific task of drafting the NIH budget proposal; in the Senate, the Committee on Health, Education, Labor and Pensions is in charge of providing recommendations for the NIH budget. Table 1 presents the list of groups I consider in this study and some information on their role and their size.

2.2 Control through the NIH organization

The partition of the NIH itself results from congressional approval, a bill can create or terminate an institute. A bill may also create other forms of autonomous programs, such as centers, to direct funds to any specific area or interest. Between 1950 and 2002, the number of institutes composing the NIH thus increased from 2 to 26. Table 12 reports the list of institutes in 2002. Overall, there is thus little institutional limitation on congressmen and the president's formal power over the distribution of public funds across research topics.

3 Description of the Data

I use two sets of data for this study: data on mortality by districts or states represented in the congressional groups reported in table 1, and data on the NIH type R grants. Both sets of data span the period 1985-2002, i.e. nine terms. In addition to the two previous sources of data, I extracted the composition of every congressional group from the Congressional directories.

The lower limit of the period of study stems from the difficulty to have a reliable estimate of mortality due to HIV/AIDS or of medical research on that disease before 1985. The appearance and the quick rise of HIV in the eighties as a major cause of deaths and as a main topic of research induced substantial changes in the recording of diseases and classification of research topics. The upper limit on the period of study stems from the unavailability, to my knowledge, of demographic data at the district level after the redistricting of 2003 (see Adler 2008, Lublin 1997). Summary statistics for the main variables defined below are reported in Table 2.

3.1 Data on mortality and congressional groups

Data on mortality stem from the *Vital Statistics* files, which contain information from deaths records for each year of the period, at the individual level. For every deceased person i, these data files report the main cause of death, the age of the deceased person, denoted A_i, the state of the deceased, and, depending on the year of death, additional demographic information. Given the set $\Omega(c, d, s, y)$ of residents state s, who died of disease d in year y, and share some additional demographic characteristics c, I define the variable $m(c, d, s, y)$:

$$m(c, d, s, y) \equiv \sum_{i \in \Omega(C,d,s,y)} (100 - A_i) \tag{1}$$

$m(c, d, s, y)$ is the *true* sum of life-years lost to 100 years old among the residents of s sharing the characteristics c, in the year y, because of disease d.

For a given senator S (respectively a representative R), let $\Omega(S, d, s, y)$ (respectively $\Omega(R, d, s, y)$) be the set of deceased individuals who were residents of a state represented by senator S (respectively of a district represented by representative R at the House). For any senator representing s, the variable $m(S, d, s, y)$ is thus the sum of life-years lost due to d in the whole state s.

With the previous notations, I denote $m^*(c, d, s, y)$ the *estimated* value of $m(c, d, s, y)$ using available data. The value $m^*(S, d, s, y)$ is computed using equation 1 and the available data from *Vital statistics*. It may differ from the actual $m(S, d, s, y)$ if some deaths are not recorded, or information on the deceased individual's state of residence is missing, etc.

Since there is no information on deaths at the district level in the *Vital Statistics* files, $m^*(R, d, s, y)$ cannot be estimated directly. Instead, I use Adler 2008 and Lublin 1997, who provide aggregated data by district and by Congress term, to compute $m^*(R, ., ., .)$. Let $w(R, c, s, t)$ the share of individuals with characteristics c in s in Congress term t who reside in the district that R represents and C the set of demographic characteristics that partition the set of individuals. I compute $m^*(R, d, s, y)$ as:

$$m^*(R, d, s, y) = \sum_{c \in C} w(R, c, s, t) m^*(c, d, s, y) \tag{2}$$

where $m^*(c, d, s, y)$ is the estimation of equation 1 with the *Vital statistics* data. By definition, $\sum_{R, representative\ of\ s} \sum_{c \in C} w(R, c, s, t) = 1$. In addition to the errors stemming from the construction

of the data, the variable $m^*(R, d, s, y)$ will differ from the actual $m(R, d, s, y)$ if mortality conditions are heterogeneous in the population sharing characteristics c in the state s. This measurement may then affect the estimations of the next sections. To check their robustness, I use three different sets of demographic characteristics in the equation 2 that correspond to three partitions of the set of deceased individuals:

- set of characteristics $C_1 \equiv \{urban, rural\}$,
- set of characteristics $C_2 \equiv \{government\ employee, other\}$,
- set of characteristics $C_3 \equiv \{black, white, other\}$.

These partitions are aimed to match information on deceased individuals from the *Vital statistics*, and information on district composition used to compute $w(R, c, s, y)$ variables, which are provided in Lublin 1997 or Adler 2008.[7]

For any group g of n_t^g congressmen in Congress term t, I define the normalized life-years lost function LY as:

$$LY_{d,t}^g \equiv \frac{1}{100 \times n^g} \sum_{X \in g \ at \ t} m^*(X, d, s, y) \tag{3}$$

The variable $LY_{d,t}^g$ represents the average sum of life-years lost because of disease d in year y, in the district (resp. state) of the representatives (resp. senators) X belonging to the group g in Congress term t (counting twice a state if both its senators belong to g).

Remarks: the previous equations allow the year variable y of the mortality files to differ from the Congress term variable t of the congressional groups and congressional districts data. In what follows, I use the *Vital Statistics* file for the year 1989 to estimate life-years lost variables for all the groups of Table 1 and all the Congress terms, that is I compute life-years lost as though the distribution of mortality had been constant throughout years, equal to mortality in 1989.[8] I use mortality data *from a single year*, instead of data of the current years corresponding to the Congress term t, to avoid potential sources of biases when estimating the impact of life-years lost on research spending. If some

[7] Since Lublin 1997 and Adler 2008 do not contain information after 1998, I assume that the composition of the districts did not change between 1998 and 2002 to compute life-years lost in districts after 1998. The redistricting of 2003 prevents to interpolate district data after 2002.

[8] Districts composition or borders may vary throughout terms, though.

investment in medical research on a disease in a given year decreases mortality due to this disease the following years, changes in mortality from one year to another will be highly correlated across groups, assuming groups are not too different from the US population. This correlation will create collinearity across the first-difference of life years lost variables, which could bias the estimation of the coefficients. In fact, since the change in mortality in any group would then be correlated with virtually *any* sample of US residents that is close enough to the whole population, the estimation would suffer from omitted-variable bias.

In practice, using mortality data from 1989 or from some other year shouldn't change the results of the estimations, since the distribution of mortality across states and diseases has been extremely stable over time.[9] However, I chose 1989 because Vital statistics before that year do not provide as detailed information on the ethnicity of the deceased, which I use to compute life-years lost at the district level.

Table 13 reports summary statistics of life-years lost by disease, for the House as a whole in 1989, and for the House majority, and the House Appropriations Subcommittee for the NIH. The set of thirteen diseases (or disease categories) retained in this study causes more than 90 per cent of the deaths in the US in 1989. All the other deaths are due to accidents, murders, or unidentifiable causes. The table shows the correlation between the distribution of life-years lost across diseases in these groups. The standard deviation of life-years lost in the House majority and the House Subcommittee reflect changes in the composition of these groups every Congress term, a necessary condition for the empirical strategy.

3.2 Data on NIH grants

The second set of data derives from detailed information on R01 grants at the grant level provided in various NIH databases: the *REsearch Portfolio Online Reporting Tools (REPORT)* and the NIH data supplied by the NBER (Lichtenberg 2001) for the period 1986-1991.[10] I use these data to gather the following information for every competing "type R" grant: the fiscal year the grant was awarded

[9] I checked that this claim is true using mortality data for 1985, when possible, and 1998.

[10] For a previous version of this paper, I used a subsample of these data that were available on the *Computer Retrieval of Information on Scientific Projects* (CRISP) database, a website that has been discontinued and replaced by the *REPORT* database. Some of the estimations differ quantitatively, but not qualitatively from the results presented here. Except for the NBER data, all the other data are now available on http://report.nih.gov/ or from http://projectreporter.nih.gov/reporter.cfm.

(which is by definition one year after the budget is voted on), the total amount of funds received by the grantee for the whole length of the grant, the state where the research institution of the grantee is located, the specific agency within the NIH awarding the grant and a set of keywords describing the research project funded by the grant. "Type R", and especially *R*01, is by far the most common type of grant. The NIH documentation, available on http://grants.nih.gov/grants/funding/r01.htm, mentions: "The Research Project Grant (R01) is the original and historically oldest grant mechanism used by NIH. The R01 provides support for health-related research and development based on the mission of the NIH. [...] The Research Project (*R*01) grant is an award made to support a discrete, specified, circumscribed project to be performed by the named investigator(s) in an area representing the investigator's specific interest and competencies, based on the mission of the NIH." Matching other grants to disease categories is much harder, since they either cover too many different topics, or no specific topic.

The keywords describing a grant mention whether the research hereby funded involved human subjects. I define such research as *clinical*, and the other projects as *basic*.

When possible, I also match every grant to a unique disease in the set of disease or disease categories retained here (see Table 13). The matching is not possible for some grants that support research on sets of conditions that are too large to be matched with a unique disease (e.g. *pathology*), or that are related to non-lethal conditions (e.g. *back injury*).[11] Finally, due to missing information, some grants cannot be matched to a US state.

For the purpose of this paper, I aggregated these data so as to obtain the total amount of funds awarded by disease, by type of research and by Congress term (and by state, for some specifications). I also use series of appropriations aggregated by institute (available on the http://www.nih.gov/about/almanac/appr for some additional results.

Remark. Table 13 reports some summary statistics by disease and type of research. They show the positive correlation between the funds for research on a disease, either basic or clinical, and the share of life-years lost, in the US or in the House majority. The major change in life-years lost occurred in 1995 (after the 1994 elections), when the majority of the House became Republican, in figure 1, I plot the net increase in NIH type R funds with respect to the net increase in the average log of life-years lost in House majority districts before and after the 1994 elections. The graph shows a positive

[11] The matching strategy follows Toole 2007.

correlation. The purpose of the estimations of the next sections is to refine this result.

4 Empirical Strategy

4.1 Specification

Consider a group g of interest, $LY_{d,t}$ the share of life-years lost due to disease d in the districts represented in the group g at t, formally defined in equation 3.[12] The objective of the empirical part is to estimate the effect of these two variables on the allocation of public funds for medical research. It relies on the specification:

$$\log F_{d,\theta,t} = c_{d,\theta} + \beta \log LY_{d,t} + \beta_{Clinic} \log LY_{d,t} \times Clinic_{\theta} + \mu X_{d,\theta,t} + \epsilon_{d,\theta,t} \tag{4}$$

where $F_{d,\theta,t}$ is the amount of funds for research on disease d of type θ awarded in Congress term t, and $Clinic_{\theta}$ is a dummy variable equal to one if $\theta = clinical$, and 0 if $\theta = basic$. The vector of variables $X_{d,\theta,t}$ may include control variables such as life-years lost variables for other groups, as well as Congress term dummy variables, and their interactions with $Clinic_{\theta}$. The term $c_{d,\theta}$ is the unobserved fixed effect for research on disease d of type θ. In all estimations, standard errors are clustered at the disease and type of research level.

For some additional results, I also use a finest level of aggregation to estimate a linear model of the form:

$$\log F_{d,\theta,s,t} = c_{d,\theta,s} + \beta \log LY_{d,t} + \beta_{Clinic} \log LY_{d,t} \times Clinic_{\theta} + \rho N_{s,t} + \rho_{Clinic} N_{s,t} \times Clinic_{\theta} + \mu X_{d,\theta,s,t} + \epsilon_{d,\theta,s,t} \tag{5}$$

where $N_{s,t}$ is the number of congressmen in group g who represent the state s, or a district of that state, and where $F_{d,\theta,s,t}$ is the amount of funds for research on disease d of type θ awarded to research institutions in state s, in Congress term t. The vector of variables $X_{d,\theta,s,t}$ may now include control variables such as life-years or number of state congressmen variables for other groups, as well as Congress term dummy variables, and their interactions with $Clinic_{\theta}$. The term $c_{d,\theta,s}$ is the unobserved

[12] The superscript $.^{g}$ on these variables, displayed in equation 3, is omitted here for readability.

fixed effect for research on disease d of type θ performed in state s. In all estimations, standard errors are clustered at the disease and type of research and state level.

4.2 Identification assumptions

The first objective of this paper is to identify non-scientific factors affecting the funding of publicly-funded medical research. The factors studied here, life-years lost across diseases in some congressional group, result from the recomposition of that group following Congressional elections, every two years. Hence, the identification holds if the selection of members of that group is not based on a systematically higher (or systematically lower) productivity of research on the diseases affecting disproportionately their constituents (Assumption 1).

The coefficients β and β_{Clinic} also provide information on the balance of powers in Congress under the additional identification assumption that, on average, members of a given group *do not derive a larger utility from medical research funding* than other congressmen (Assumption 2). In other words, members of that group are not "preference outliers". If this assumption fails, a positive coefficient β (or β_{Clinic}) may just stem from the fact that the NIH budget reflects current Congressmen's preferences. If this assumption holds, a positive β would indicate that members of that group have a larger bargaining power on the allocation of NIH funds, a larger bargaining power they obtain through their membership to that group.

Do Assumptions 1 and/or 2 hold? Let's consider majority groups of either chamber first. The affiliation of a congressman to the majority party is a direct consequence of congressional elections. If they are elected, members of any party may a priori take part in any debate regarding the budget, not only the NIH budget. Since the timing of elections is exogenous, and medical research priorities have never been a theme of congressional electoral debates, the recomposition of these groups should be independent of medical research concerns. Assumptions 1 and 2 should thus hold for these groups.[13]

Let's now consider the Appropriations committees and subcommittees. Frisch and Kelly 2006 or Schneider 2007 describe the procedures leading to the composition of these groups: following each election, the distribution of committee affiliations is decided within each party, after each congressman has submitted the list of committees, ranked by order of preferences. The distribution of subcommit-

[13] A few congressmen switched parties at some point of the period I study. Counting them in either party does not change the following estimations.

tee seats follows a similar procedure. No congressional rule imposes any constraint on the matching between preferences and groups memberships and, in practice, a congressman's assignment to a sub/committee will depend on her/his preferences and clout at the time of the process. In particular, no feature of the membership process indicates that Assumption 1 fails for any of the sub/committees I consider here, and I could not find any congressional document, or any feature of the affiliation process, suggesting that the composition of these groups were linked to research productivity factors.

Given that sub/committees members have specific areas of specialization, Assumption 2 may be more questionable for these groups. Members of both subcommittees (in the House and in the Senate), however - and even more so the whole Appropriations committees - are in charge of several types of spending besides the NIH budget, such as education or labor issues. In addition, congressional rules limit the number of committees and subcommittees a congressman may belong to, as well as the number of members of Appropriations subcommittees. In addition, the total discretionary budget that any budget authority agency will be in charge of seems mostly exogenous, heterogeneous across agencies, and roughly stable over the period studied from the information on discretionary spending provided in Tables 5.1 to 5.6 available on http://www.whitehouse.gov/omb/budget/Historicals/. These three last facts combined suggest that any congressman can assess the expected share of federal funds that membership to any subcommittee he or she will have authority over, and may rank subcommittees according to that share.[14] The variety of issues falling under the responsibility of the members of HouS and its Senate counterpart, and the heterogeneity of budget sets across Appropriations subcommittees thus support Assumption 2 for these groups. To my knowledge, no study of the selection of congressmen into the subcommittees indicates that its members have any distinct interest for medical research priorities. (As a robustness check though, I perform an instrumental variable analysis to relax Assumption 2 for HouS. The instrument is contingent to the results of the estimations so that, to simplify the presentation, I report the discussion of this approach to the next section).

Some other factors, inherent to the construction of the life-years lost variables, may also bias the estimations and were discussed in the section 3.1. I will use the set of characteristics $C3$ of section 3.1 to compute life-years lost at the district level for the first estimations in the next section, then

[14] In spite of the increase of the NIH budget in the late nineties the distribution of these shares have been roughly stable over the period (see Dalton 2000 for information on the increase and "Table 5.5. Percentage Distribution of Discretionary Budget Authority by Agency: 1976-2015" on http://www.whitehouse.gov/omb/budget/Historicals/ for information on the budget).

perform these same estimations using the two other possible life-years lost.

5 Results

5.1 Allocation of funds across diseases

This section focuses on the impact of life-years lost across diseases and groups on the allocation of funds across diseases, i.e. on the estimation of the coefficients β and β_{Clinic} of equations 4 for the set of groups of table 1.

Regressions at the disease, type of research level. Table 3 reports the OLS estimation of the coefficients of the first difference of equation 4, separately for every group, twice.[15] The first series of estimations shows that life-years lost in three groups may impact NIH funds: HouS, its equivalent in the Senate, and the House majority. In all three cases, the effect is larger clinical than for basic research. In fact, the effect of HouS variables on basic research is not significantly different from zero here. The second series of estimations (columns 8 to 13), indicate that the effect of HouS holds when controlling for other groups. Depending on whether we take into account the estimated coefficient $\hat{\beta}$, a 1 percent increase in life-years lost in districts represented in HouS leads to a 0.8-1.1 percent increase in funds for clinical research.

Regressions at the disease, type of research and state level. The previous specifications fit best the actual formulation of the budget described in section 2. However, as a robustness check, I estimate the parameters of equation 5 in Table 4, with robust standard errors clustered by disease, type of research and state.[16] These estimations confirm the effect of life-years lost in districts of HouS members on clinical research and the absence of effect of this group on the funds for basic research. This effect is robust to the inclusion of variables $N_{s,t}$, which control for the potential earmarking of research funds to the representatives' states. (The estimated coefficients of $N_{s,t}$ are reported later for clarity.) The information provided by this regression suffers from the fact that the amount of research across states is extremely heterogeneous, however. This heterogeneity likely explains the negative coefficient of life-years lost in the House Appropriations committee. In fact, the effect of any group, other

[15] The small size of the sample and the correlation between life-years lost variables would limit the interpretation of a regression including all life-years lost variables at once.

[16] As mentioned in the Section 3, I assume that the missing observations (less than 200 out of 10400) result from missing information on the amount of funds for research on a given disease of a given type in a given state. A Tobit estimation would give similar results.

than HouS, disappears when I use the level of funds (in $ millions), as the dependent variable, and average life-years lost as defined in equation 3, and not the Log of these variables, in the estimation of equation 5 (Table 5). Given that HouS contains around 14 members, and that data are aggregated by Congress terms (that is two years), columns 1,2 or 3 in Table 5 show that an additional life-year lost due to a disease in HouS leads to an increase of NIH funds by year for clinical research on that disease equal to around $\frac{6.4\times10^6}{2\times14\times100} \simeq$ \$2300 by year.

Measurement error of life-years lost variables. Table 6 reports the result of the regression in first differences of variants of equation 4 for the three different estimations of life-years lost at the district level defined by the set of characteristics $C1$, $C2$ or $C3$, used to partition Vital statistics data (see section 3.1). For ease of comparison, I report the results for all three possible measures of life-years lost (so that column (9) in Table 6 and column (5) of Table 3 are the same).

The estimations of columns with most controls show that the effect of HouS is more robust to changes in the method used to compute life-years lost than the effect of House majority, although life-years lost in the House majority may have both an impact on basic and clinical research. Overall, all estimations are consistent with the previous results. No coefficient is significantly negative. These results support the hypothesis that measurement error induced by the lack of mortality data at the district level biases the effect of life-years lost in either group on the NIH budget towards zero.

Whose bargaining power is it? If Assumption 2 does not hold for HouS congressmen, interpreting the impact of life-years lost in HouS as evidence of bargaining power of its members is questionable. Assumption 2 more likely holds for the composition of the House majority, which also has an effect on NIH funds. Both variables may be correlated, however, since congressional rules prescribe the share of majority members in HouS to approximate the share of majority members in the whole House.

The correlations reported in columns (1), (4) and (7) of Table 7, controlling for Congress terms, show that it is indeed the case. In fact, the more correlated they are, the more significant the effect of House majority life-years lost variable is. For completeness, I also report the second stage of the SLS estimation that uses the first difference of life-years lost in House majority districts as an instrument for life-years lost in HouS districts (although columns (2),(5) and (8) are essentially the same as columns (2),(5) and (8) of Table 6).

To summarize, life-years lost in the districts represented in HouS districts have a significant and robust impact on the allocation of funds across diseases, more so for clinical than for basic research. This impact stems from some additional bargaining power congressmen obtain from belonging to either HouS, the majority of the House or any of these groups.

Counterfactual estimation Which diseases benefited from the 1995 majority change? To address this question, I use the estimated coefficients of Column (4) in Table 3 to estimate the average yearly allocation of funds across diseases after 1994 if majority had not changed in 1994. The results are presented in Table 8. I find that the majority change mainly affected negatively funding on HIV and Congenital disorders, and affected positively funding for research on Central Nervous System disorders, Psychiatric conditions, and Respiratory disorders.

5.2 Allocation of funds across states

Table 9 reports the estimations of the coefficients ρ and ρ_{Clinic} of equation 5 and some variants. The number of representatives of a state in HouS shows no significant positive effect on the research funds for that state, regardless of whether other groups' variables are included in the regressions. (In fact, it seems to have a negative impact on the funds by state; this result, however, is not robust to the inclusion of additional controls). No group seems to have any significant positive effect on the distribution of funds across states. These results do not imply that earmarking does not exist at all — in fact some empirical studies and anecdotal evidence suggest the opposite — but they suggest that pork-barrel either affects the redistribution within states or involves relatively small amounts of funds.

The lack of any significant effect is quite surprising nevertheless, especially for the specifications that do not include the social cost variables (e.g. Column 1 of table 9), since one may conjecture that the research interests of research institutions are correlated with the health conditions of their neighboring population. To investigate further that conjecture, I compute the log of the share of life-years lost $LY^s_{d,t} \equiv \frac{life-years^s_d}{\sum_d life-years^s_d}$ separately for every state, and regress the log of funds on that variable, as in:

$$\log F_{d,\theta,s,t} = \omega \log LY^s_{d,t} + \omega_{Clinic} \log LY^s_{d,t} \times Clinic_\theta + X_{d,\theta,s,t} + u_{d,\theta,s,t} \tag{6}$$

where $X_{d,\theta,s,t}$ may include dummy variables for every Congress term, type of research, disease and

state, and the interactions between the type of research and other dummy variables. The results (see Table 10) do show some significant impact of life-years lost on research funds at the state level, but this impact is not significant once additional control variables are included. In fact, the figure 2, which represents the normalized log of funds with respect to the normalized log of life-years lost for every state shows that no correlation is visible for most diseases.

6 Discussion

6.1 What causes a group's bargaining power?

As was demonstrated in the seminal paper of Baron and Ferejohn 1989, and confirmed in the empirical work of Knight 2005, being the proposer provides bargaining power, if congressmen are impatient, when the issue at stake is the distribution of some budget across states or districts. Among the groups I consider, the actual proposer is indeed HouS. Its Senate equivalent also participates to the design of the proposal, but its members by definition do not participate to the initial vote of the budget, which occurs in the House. Similarly, the composition of the Senate Appropriations subcommittee on Labor and Health can make recommendations, which are non-binding, and do not seem to steer funds towards diseases affecting more their constituents.

In addition to these theoretical arguments, the quotation from Varmus 2009 also suggests that HouS members may tend more to contact directly the NIH staff, whom they should meet more often than the rest of Congress due to their role in the Appropriations process. Apart from access to NIH staff, HouS members may also have a better understanding of how the allocation of funds across institutes or autonomous programs will impact actual funding by disease.

6.2 Why is the effect on clinical research larger?

By definition, clinical research is more applied and, as such, should lead faster to pharmaceutical innovations. In addition, the probability that some investment in research on a given disease indeed leads to some discovery on this disease, and not on another one, may be larger for clinical than basic research. However, these arguments could explain why congressmen prefer clinical to basic research, not why they seem less prone to influence the allocation of funds for basic research than for clinical research.

The lack of influence of congressmen could instead stem from the fact that the productivity of clinical research may be more easily estimated by non-specialists of medical research than basic research. This difference derives from the complexity of scientific knowledge as well as from the legal restrictions that bind the uncertainty of clinical research, which, by definition, uses human data. If congressmen' s information on basic research is poor enough, they should delegate the allocation of these funds to the NIH staff, who shares congressmen's concern for research productivity, but has no reason to take into account the specific health conditions of powerful congressmen's electorate.[17] I develop this hypothesis in a model presented in Appendix. An extended version of this model is the basis for the specification of the empirical part.

Although the model is substantially different from Aghion and Tirole 1997, to fit the specific question of this paper, proposition 1 raises a similar point: congressmen's authority on the distribution of funds for basic research is *formal*. Due to their lack of information on the productivity of this type of research, they prefer to delegate the allocation decision to the NIH. Anecdotal evidence from the description of the NIH organization also supports that hypothesis: the explicit mission of one agency, the *National Institute of General Medical Sciences*, is to fund basic research on any disease. The allocation of the budget of this agency across diseases is thus effectively delegated to its staff.[18] As argued by Foucault 1963, political authority influences the focus of medical research, at least through the funding of this research. However, the concern for scientific productivity limits the effect of this authority on basic research.

6.3 Channel of congressmen's influence

According to the rules of the budget procedure, the unique channel of congressmen's influence lies in their authority over the allocation of funds across the NIH autonomous units (in addition to the possibility to create such units, which rarely occur in comparison to the yearly Appropriations process). If their preferences indeed depend on the distribution of funds across diseases (and type of research), they then have only the possibility to maximize their *expected* utility knowing the distribution of funds across those autonomous units. To investigate this point, I replace $\log F_{d,\theta,t}$ in equation 4 with $E[\log F_{d,\theta,t}|\{F_{i,t}\}_{i\in NIH}]$, where $\{F_{i,t}\}_{i\in NIH}$ is the distribution of total budgets awarded to every

[17] NIH officials are usually appointed for several years.

[18] Congressional bills may yet give some directions for the use of that budget. Moreover, a substantial number of basic grants are awarded by other agencies or smaller subdivisions of the NIH.

autonomous unit of the NIH.[19]

I use two methods to derive $E[logF_{d,\theta,t}|\{F_{i,t}\}_{i\in NIH}]$. The first method consists in running the regression:

$$\log F_{d,\theta,t} = \sum_{i\in NIH} \alpha_{d,\theta,i} \log F_{i,t} + \gamma X_{d,\theta,t} + \epsilon_{d,\theta,t} \tag{7}$$

where $F_{i,t}$ is the amount of funds allocated to institute i of the NIH at t, and $X_{d,\theta,t}$ potential additional controls, and then use the predicted value $\log \hat{F}_{d,\theta,t}$ as an estimation of $E[\log F_{d,\theta,t}|\{F_{i,t}\}_{i\in NIH}]$. Since the size of the sample is small in comparison to the number of NIH units, the predicted value will be so close to the actual value of $\log F_{d,\theta,t}$ that the results of the estimation of the coefficients of 4 will be essentially the same as in Table 6.[20]

The other method is to compute first the average share of funds by disease and by type of research for every institute $\beta_{d,\theta,i,t} \equiv \frac{\sum_{y<t} F_{d,\theta,i,y}}{\sum_{y<t} F_{i,y}}$ composing the NIH *for the years preceding the Congress term* t, where $F_{d,\theta,i,y}$ is the amount of R grants awarded in year y to research on disease d and type θ by the institute i. I then estimate $E[\log F_{d,\theta,t}|\{F_{i,t}\}_{i\in NIH}]$ as $\sum_{i\in NIH} \beta_{d,\theta,i} \log F_{i,t}$, that is I approximate the expected log of funds with the log of expected funds. The results of the estimation of equation 4 using this latter method are shown in Table 11. Although they should be interpreted in regard of the lack of any obvious way to know how congressmen form their expectations about NIH funds, these estimations strongly support the hypothesis that the House majority members try to steer NIH funds towards diseases that affect more their constituents through the possibility to direct funds towards NIH autonomous units, whereas the smaller (yet significant) coefficients of HouS life-years lost suggest that HouS members may have more effective ways to influence the NIH budget than this channel.

7 Conclusion

This paper finds that the social cost of a disease, measured in life-years lost, on the constituents represented either in the majority party of the House or in the subcommittee of the House Appropriations Committee in charge of health issues has a positive impact on the NIH funds for clinical research

[19] These budgets thus include both funds that will be awarded through type R grants, and other types of spending.

[20] In fact, the model is even over-identified if I estimate 7 aggregating observations at the Congress term level. Instead, I thus estimate it aggregating NIH funds at the year level.

on that disease awarded through R grants. In addition, there is no significant effect of the state origin of the members of these groups on the allocation of funds across states.

The main channel of influence of congressmen over the medical research agenda relies on their control of the allocation of funds across the multiple units composing the NIH. This paper thus shows that the change in the rules that govern the *rules of formulation* of the budget would limit congressmen's influence. A more radical approach would be to reduce the number of these units. This change is advocated by Varmus 2009, who "expressed [his] anxieties about the continuing proliferation of autonomous units at the NIH".[21] Such a change may decrease congressmen's incentives to devote such a large of federal funds to support medical research, however. Whether such a change would benefit scientific research, or have an overall positive impact on welfare, yet remains an open question.

[21] p.174

References

Acemoglu, Daron (2002), "Directed Technical Change," *Review of Economic Studies*, LXIX, 781-810.

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.

Acemoglu, Daron, David Cutler, Amy Finkelstein, and Joshua Linn, (2006) "Did Medicare Induce Pharmaceutical Innovation?", *American Economic Review*, 96(2): 103-107.

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 Peter Howitt (1992), "A Model of Growth through Creative Destruction," *Econometrica*, LX, 323-351.

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

Arrow, Kenneth (1962), "Economic Welfare and the Allocation of Resources for Invention." in Nelson, R. (Ed.), *The Rate and Direction of Inventive Activity.* Princeton, NJ.

Atlas, Cary, Thomas Gilligan, Robert Hendershott, and Mark Zupan "Slicing the federal government net spending pie: who wins, who loses , and why", *American Economic Review,* 85.

Atwood, K.C. , E. Woeckner, R.S. Baratz, W.I. Sampson "Why the NIH trial to assess chelation therapy (TACT) should be abandoned. *Medscape J Med* 2008;10(5):115.

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

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

Bhattacharya, Jayanta and Mikko Packalen (2008), "Is Medicine an Ivory Tower? Induced Innovation, Technological Opportunity, and For-Profit vs. Non-Profit Innovation", NBER working paper #13862.

Blume-Kohout, Margaret E., Krishna B. Kumar and Neeraj Sood (2009), "Federal Life Sciences Funding and University R&D", NBER working paper #15146.

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

Cohen, Linda and Roger Noll (1991), *The Technology Pork Barrel*, Brookings Institution Press.

Cutler, David, Ellen Meara and Seth Richards (2009), "Induced Innovation and Social inequality: Evidence from Infant Medical Care", NBER working paper #15316.

Dalton, Rex (2000), "Clinton proposes $2.8 billion increase in science funding" *Nature* 403, 349 (27 January 2000)

David, Paul A., Bronwyn H. Hall and Andrew A. Toole (2000), "Is public R&D a complement or substitute for private R&D? A review of the econometric evidence", *Research Policy* 29, 497-529.

Drandakis, Emmanuel and Edmund Phelps (1966), "A Model of Induced Invention, Growth and Distribution", *The Economic Journal*, December.

Fisher Ellison, Sara and Catherine Wolfram (2001), "Pharmaceutical prices and Political Activity", NBER working paper #8482.

Ferejohn, John (1974), *Pork Barrel Politics; River and Harbors Legislation, 1947-1968*, Stanford University Press.

Finkelstein, Amy (2004), "Static and Dynamic Effects of Health Policy: Evidence from the Vaccine Industry." *Quarterly Journal of Economics,* 119(2), pp. 527-64.

Foucault, Michel (1963) [1994], *The Birth of the Clinic: An Archaeology of Medical Perception*, Vintage.

Frisch, Scott and Sean Kelly (2006), *Committee Assignment Politics in the U.S. House of Representatives* Oklahoma University Press.

Griliches, Zvi (1991), "The Search for R&D Spillovers". NBER working paper #3768.

Griliches, Zvi, Tor Jakob Klette and Jarle Møen (2000)"Do Subsidies to Commercial R&D Reduce Market Failures? Microeconomic Evaluation Studies" *Research Policy* Volume 29, Issues 4-5, Pages 471-495.

Grossman, Gene, and Elhanan Helpman (1991), *Innovation and Growth in the Global Economy*. Cambridge, MA: MIT Press.

Hall, Bronwyn H., Jacques Mairesse and Pierre Mohnen (2010), "Measuring the returns to R&D", in *Elsevier Handbook of the Economics of Innovation*, B. H. Hall and N. Rosenberg (eds.), April 2010.

Hayami, Yujiro and Vernon Ruttan (1970), "Factor prices and technical change in agricultural development: the US and Japan, 1880-1960," *Journal of Political Economy*, LXXVII, 1115-1141.

Hegde, Deepak (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.

Jacob, Brian and Lars Lefgren (2007), "The Impact of Research Grant Funding on Scientific Productivity" NBER working paper #13519

Jayachandran, Seema (2007) "The Jeffords Effect", *The Journal of Law and Economics* UChicago Press.

Knight, Brian (2005), "Estimating the Value of Proposal Power", *American Economic Review,* 95(5), 1639-1652.

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

Levitt, Steven and James Snyder (1997) "The impact of federal spending on house election outcomes." *Journal of Political Economy*, 105.

Lichtenberg, Frank R. (2001), "The Allocation of Publicly-funded Biomedical Research." Chapter 15 in David M. Cutler and Ernst R. Berndt, eds., Medical Care Output and Productivity, Chicago: University of Chicago Press, pp. 565-589, 2001.

Lichtenberg, Frank R. and Joel Waldfogel (2009) "Does Misery Love Company? Evidence from pharmaceutical markets before and after the Orphan Drug Act," 15 Michigan Telecommunications and Technology Law Review, 335.

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

David Lublin. 1997. "Congressional District Demographic and Political Data," American University, Washington, D.C.

Miller, Grant (2008), "Women's Suffrage, Political Responsiveness, and Child Survival in American History," *Quarterly Journal of Economics* 123(3): 1287-1327.

Murphy, Kevin M. and Robert Topel (2003), "The Value of Health and Longevity" *Journal of Political Economy* vol. 114, No. 5.

arrison Nelson, Committees in the U.S. Congress, 1947-1992

Payne, Abigail (1999), "Are All Agencies Alike? The Impact of Politics on University Research Funding From Ten Federal Agencies," mimeo.

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.

Romer, Paul M. (1990), "Endogenous Technological Change," *Journal of Political Economy*, XCVIII, S71-S102.

Samuelson, Paul (1965), "An induced innovation along Kennedy-Weisacker lines", *Review of Economics and Satistics*, XLVII, 444-464.

Schneider, Judy (2007), " House Subcommittees: Assignment Process," *CRS Report for Congress* 98-610 GOV.

Snowberg, Erik, Justin Wolfers, and Eric Zitzewitz (2006), "Party Influence in Congress and the Economy" NBER Working Paper No. 12751.

Charles Stewart III and Jonathan Woon. Congressional Committee Assignments, 103rd to 111th Congresses, 1993–2009

Toole, Andrew A. (2007), "Does Public Scientific Research Complement Private Investment in Research and Development in the Pharmaceutical Industry?" *Journal of Law and Economics* vol. 50 p.81-103.

Varmus, Harrold (2009), *The Art and Politics of Science* W.W. Norton and Company, New York - London.

8 Figures and Tables

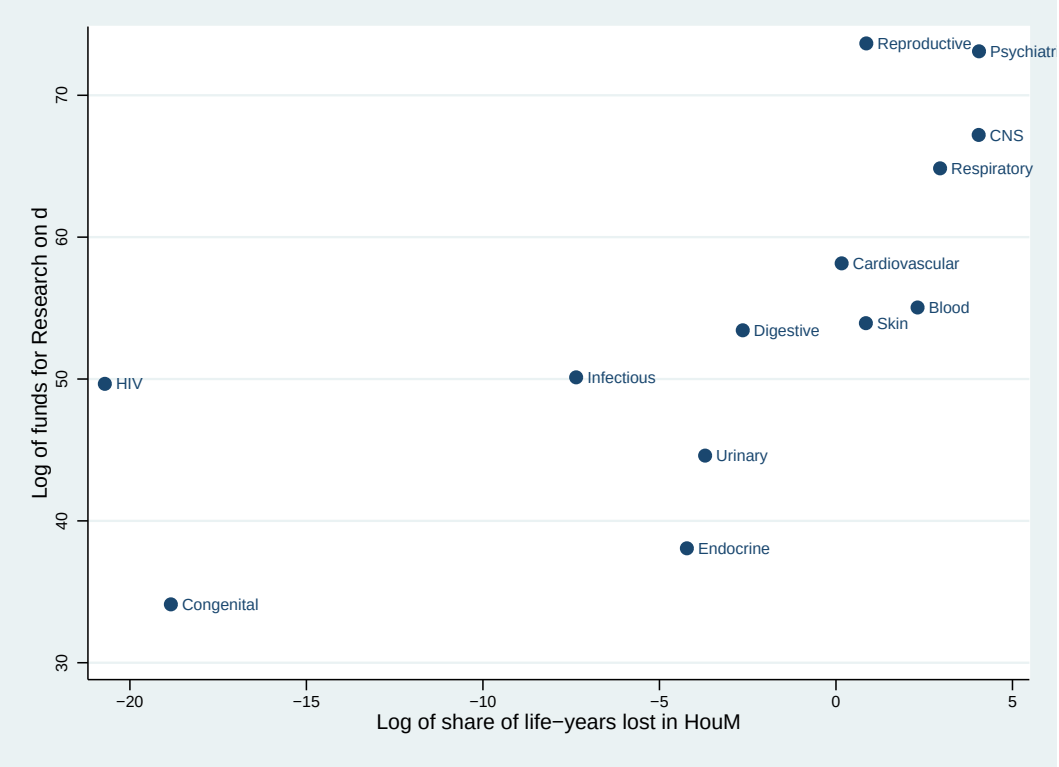

Figure 1: Increase in life-years lost in House majority districts and increase in funds for research before and after 1994.

Notes: The x-axis represents the difference between the log of the share of life-years lost in the districts represented in the House majority before and after the 1995 majority change in the House. The y-axis is the difference in log of funds. The full name of diseases is reported in Table 13.

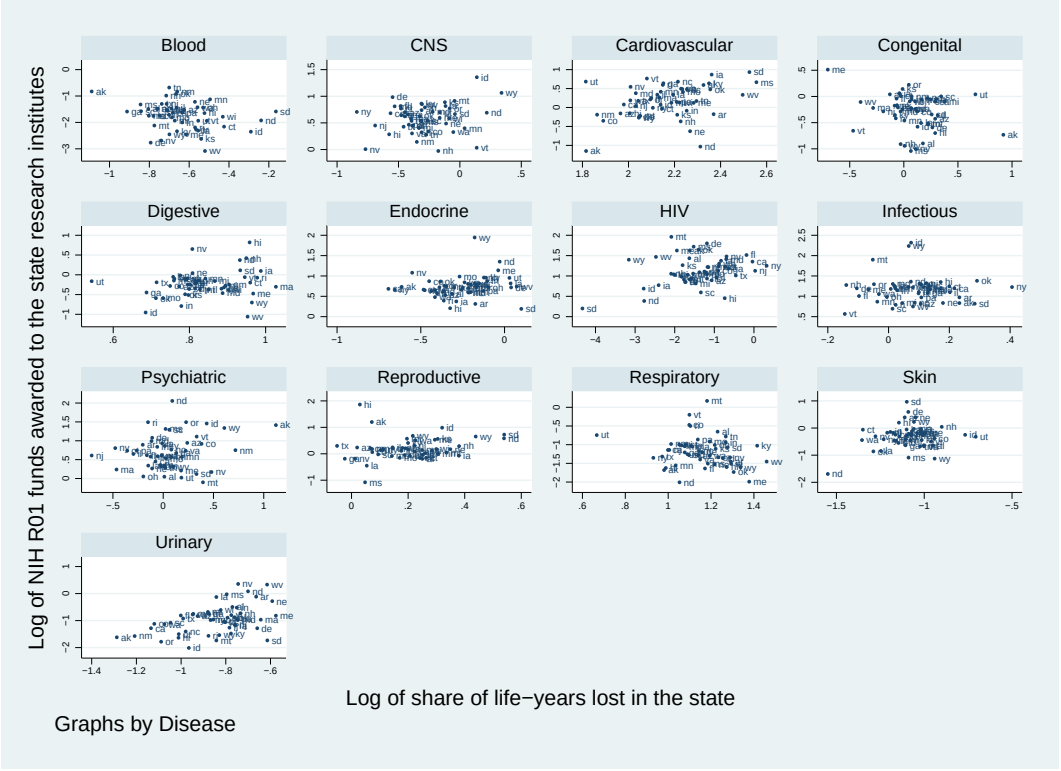

Figure 2: Share of life years lost and NIH funds awarded during the period 1985-2002, by state and by disease.

Notes: See the text for the construction of the data. The x-axis represents the log of the share of life-years lost in any state in 1991.

Table 1: **Congress Groups**

Group Name	Role	Average number of members
House Majority		245
House Appropriations Committee	designs the budget proposal	60
House Appropriations Subcommittee on Labor, Health and Human Services, Education and Related Agencies (HouS)	designs the NIH budget proposal (among other charges)	14
Senate Majority		54
Senate Appropriations Committee	designs the budget proposal	29
Senate Appropriations Subcommittee on Labor, Health and Human Services, Education and Related Agencies	designs the NIH budget proposal (among other charges)	15
Senate Committee on Health, Education, Labor and Pensions	provides recommendations for the NIH budget	17

Note: The exact name of committees and subcommittees slightly changes from year to year.

Table 2: **Summary statistics**

Variable	Mean	Std. Dev.	Min.	Max.	# Obs.
Funds $F_{d,s,\theta,t}$	10.594	24.84	0	519.036	11700
Funds $F_{d,\theta,t}$	529.695	503.419	28.343	2999	234
Log $F_{d,s,\theta,t}$	14.388	2.418	3.281	20.067	11558
Log $F_{d,\theta,t}$	19.673	0.958	17.16	21.822	234
Log of life-years lost in:					
Senate Labor Com.	566.17	756.99	68.727	3458.584	234
Senate Appropriations Com.	615.953	800.877	108.911	4069.108	234
Senate Appropriations Subcom.	581.329	775.639	96.613	4013.414	234
Senate Majority	660.639	868.89	104.237	3980.352	234
House Appropriations Com.	77.782	101.165	15.064	410.586	234
HouS	78.136	102.255	14.675	432.073	234
House Majority	77.974	101.002	15.115	408.931	234
Number of state congressmen in:					
Senate Labor Com.	0.342	0.498	0	2	450
Senate Appropriations Com.	0.576	0.55	0	2	450
Senate Appropriations Subcom.	0.3	0.482	0	2	450
Senate Majority	1.087	0.775	0	2	450
House Appropriations Com.	1.207	1.377	0	7	450
HouS	0.273	0.516	0	2	450
House Majority	4.896	5.146	0	30	450

Notes: $F_{d,\theta,s,t}$ denotes funds for research on disease d of type θ, awarded to research institutes in state s in Congress term t. The *Senate* (or *House*) *Appropriations Subcom.* is the Appropriations subcommittee for Labor, Health etc., in charge of the NIH budget proposal. Life-years lost are normalized across groups, as in equation 3. Table 1 contains information on congressional groups.

Table 3: **Increase in life-years lost and NIH funds, by disease and type of research**

Group G	HouS	House Appropriations Committee		House Majority		Senate Appropriations Subcom.		Senate Appropriations Committee		Senate Majority		Senate Labor Committee	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)
	Dependent variable is Log of funds by disease and type of research												
Log of life-years lost in:													
group G	-0.023	-0.45	-0.60	0.49*	0.56*	-0.047	-0.051	0.19	0.23	0.20	0.21	-0.079	-0.084
	(0.12)	(0.71)	(0.88)	(0.21)	(0.20)	(0.094)	(0.11)	(0.15)	(0.17)	(0.36)	(0.37)	(0.068)	(0.073)
group G $\times Clinic$	0.96*	1.39	-0.73	0.16	-0.37	0.32*	0.14	0.26	-0.034	0.13	-0.018	-0.23+	-0.15
	(0.27)	(1.12)	(1.42)	(0.44)	(0.38)	(0.14)	(0.15)	(0.23)	(0.27)	(0.49)	(0.50)	(0.13)	(0.14)
HouS			0.080		-0.14		0.018		-0.12		-0.055		-0.049
			(0.16)		(0.11)		(0.13)		(0.15)		(0.15)		(0.13)
HouS $\times Clinic$			1.09*		1.04*		0.85*		0.98*		0.96*		0.91*
			(0.31)		(0.26)		(0.31)		(0.32)		(0.29)		(0.28)
Observations	208	208	208	208	208	208	208	208	208	208	208	208	208
R-squared	0.345	0.325	0.348	0.330	0.348	0.330	0.346	0.331	0.348	0.326	0.347	0.332	0.350

Notes: this table reports the OLS estimation of the first difference of equation 4 for the period 1985-2002. HouS is the House Appropriations Subcommittee in charge of the NIH budget proposal. The dependent variable *funds by disease and type of research* is the difference between the log of NIH funds for research of type θ on disease d awarded through R grants in Congress terms t and $t-1$ ($\log F_{d,\theta,t} - \log F_{d,\theta,t-1}$). *Life-years lost in ...* is the difference between the log of the share of life-years lost to 100 due to disease d in the districts/states represented in ... in Congress terms t and $t-1$ ($\log LY_{d,t} - \log LY_{d,t-1}$). *Clinic* is a variable equal to 1 if the type of research is clinical, 0 otherwise. All regressions include dummies for Congress terms and type of research and their interactions. Cluster-robust standard errors are reported in parentheses. * p<0.05, + p<0.1.

Table 4: **Increase in life-years lost and NIH funds, by disease, type of research and state**

	(1)	(2)	(3)	(4)	(5)
Dependent variable is Log of funds by disease, by type of research and by state					
Log of life-years lost in:					
HouS	0.40+	-0.079		0.25	0.25
	(0.23)	(0.30)		(0.33)	(0.33)
HouS ×*Clinic*		0.96*		0.85+	0.85+
		(0.45)		(0.50)	(0.50)
House Majority			0.29	0.12	0.12
			(0.42)	(0.53)	(0.53)
House Maj. ×*Clinic*			-0.0068	-0.57	-0.58
			(0.71)	(0.88)	(0.87)
House Appropriations Com.				-2.63*	-2.63*
				(1.24)	(1.24)
House Appr. Com. ×*Clinic*				-0.76	-0.77
				(2.03)	(2.03)
Senate Appropriations Subcom.				-0.084	-0.085
				(0.23)	(0.23)
Senate Appr. Subcom. ×*Clinic*				0.19	0.19
				(0.36)	(0.36)
Senate Majority				-0.21	-0.21
				(0.45)	(0.45)
Senate Maj. ×*Clinic*				-0.57	-0.57
				(0.72)	(0.72)
Senate Appropriations Com.				0.41	0.41
				(0.42)	(0.42)
Senate Appr. Com. ×*Clinic*				0.33	0.33
				(0.66)	(0.66)
Senate Labor Com.				-0.089	-0.089
				(0.19)	(0.19)
Senate Lab. Com. ×*Clinic*				-0.41	-0.41
				(0.28)	(0.28)
Observations	10,212	10,212	10,212	10,212	10,212
R^2	0.021	0.021	0.019	0.021	0.023

Notes: this table reports the OLS estimation of the first difference of equation 5 for the period 1985-2002. The dependent variable *funds by disease, type of research and state* is the difference between the log of NIH funds for research of type θ on disease d awarded through R grants to research institutes in state s in Congress terms t and $t-1$. *Log of life-years lost in ...* is the difference between the log of the share of life-years lost to 100 due to disease d in the districts/states represented in ... in Congress terms t and $t-1$. Column (4) includes the number of congressmen of state s represented in any group in the control variables. All regressions include dummies for Congress terms and type of research and their interactions. Cluster-robust standard errors are in parentheses. * p<0.05, + p<0.1.

Table 5: **Increase in life-years lost and NIH funds - in levels**

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Dependent variable is funds by disease and by type of research							
				and by state			
Life-years lost:							
HouS	-1.34		-1.17	-0.027		-0.024	-0.037
	(0.87)		(1.05)	(0.027)		(0.026)	(0.042)
HouS $\times Clinic$	6.43*		6.44*	0.13*		0.13*	0.16*
	(1.95)		(2.84)	(0.041)		(0.042)	(0.057)
House Majority		-1.68	-0.72		-0.034	-0.014	0.033
		(2.73)	(2.82)		(0.047)	(0.043)	(0.044)
House Maj. $\times Clinic$		5.16	-0.073		0.10+	-0.0030	-0.084
		(4.28)	(5.62)		(0.061)	(0.058)	(0.066)
House Appropriations Com.							0.0047
							(0.12)
House Appr. Com. $\times Clinic$							0.15
							(0.15)
Senate Appropriations Subcom.							-0.0015
							(0.0015)
Senate Appr. Subcom. $\times Clinic$							-0.0033
							(0.0022)
Senate Majority							-0.0021
							(0.0015)
Senate Maj. $\times Clinic$							0.0024
							(0.0020)
Senate Appropriations Com.							0.0022+
							(0.0014)
Senate Appr. Com. $\times Clinic$							-0.00077
							(0.0018)
Senate Labor Com.							-0.00018
							(0.0010)
Senate Lab. Com. $\times Clinic$							-0.0019
							(0.0014)
Observations	208	208	208	10,212	10,212	10,212	10,212
R^2	0.341	0.332	0.341	0.037	0.036	0.037	0.039

Notes: this table reports the OLS estimation of the first difference of equation 4 in Column 1 (resp. equation 5 in Columns 2 to 4) for the period 1985-2002 with the following modifications. The dependent variable *funds by disease, type of research* (resp. *funds by disease, type of research, and state*) is the difference between the NIH funds (in $ million) awarded for research of type θ, on disease d (resp. of type θ, on disease d, to research institutes in state s) in Congress terms t and $t-1$. *Life-years lost in ...* is the difference between the average number of life-years lost to 100 (divided by 100) due to disease d in the districts/states represented in ... in Congress terms t and $t-1$, see equation 3). All regressions include dummies for Congress terms and type of research and their interactions. Cluster-robust standard errors are in parentheses. * p<0.05, + p<0.1.

Table 6: **Life-years lost and NIH funds - using various measures of life-years lost**

Characteristics	(1) $C1$	(2) $C1$	(3) $C1$	(4) $C2$	(5) $C2$	(6) $C2$	(7) $C3$	(8) $C3$	(9) $C3$
	Dependent variable is Log of funds by disease and type of research								
Log of life-years lost in:									
HouS	0.42		-0.20	0.45		-0.17	0.46		-0.14
	(0.28)		(0.15)	(0.31)		(0.16)	(0.27)		(0.11)
HouS $\times Clinic$			1.10*			1.14*			1.04*
			(0.29)			(0.32)			(0.26)
House Majority		1.25*	0.73		2.20	0.99		0.57*	0.56*
		(0.59)	(0.59)		(1.76)	(1.26)		(0.22)	(0.20)
House Maj. $\times Clinic$			0.52			1.82			-0.37
			(0.88)			(2.75)			(0.38)
Observations	208	208	208	208	208	208	208	208	208
R-squared	0.329	0.328	0.343	0.329	0.328	0.345	0.333	0.329	0.348

Notes: this table reports the OLS estimation of the first difference of equation 4 for the period 1985-2002, separately for each estimation of the life-years lost variable at the district level. $C1$, $C2$ and $C3$ refer to the set of characteristics used to compute that variable (see section 3.1). *Clinic* is a variable equal to 1 if the type of research is clinical, 0 otherwise. All regressions include dummies for Congress terms and type of research and their interactions. Cluster-robust standard errors are reported in parentheses. * p<0.05, + p<0.1.

Table 7: **Life-years lost and NIH funds - 2SLS**

Characteristics	(1) $C1$	(2) $C1$	(3) $C1$	(4) $C2$	(5) $C2$	(6) $C2$	(7) $C3$	(8) $C3$	(9) $C3$
Dependent variable	Log of life-years lost in HouS	Log of funds disease and type of research		Log of life-years lost in HouS	Log of funds disease and type of research		Log of life-years lost in HouS	Log of funds by disease and type of research	
Log of life-years lost in:									
House Majority	0.76*			0.74*			0.51*		
	(0.051)			(0.21)			(0.12)		
HouS		1.66*	0.76		2.95	1.16		1.12*	0.96+
		(0.76)	(0.74)		(2.13)	(1.70)		(0.40)	(0.49)
HouS $\times Clinic$			1.79			3.58			0.31
			(1.13)			(3.05)			(0.77)
Observations	208	208	208	208	208	208	208	208	208
R-squared	0.455	0.277	0.284	0.371	0.132	0.098	0.269	0.311	0.318
F-stat	34.31			32.64			96.60		

Notes: this table reports the 2SLS estimation of the first difference of equation 4 for the period 1985-2002, using life-years lost in the House majority districts as an instrument for life-years lost in HouS, separately for each estimation of the life-years lost variable at the district level. $C1$, $C2$ and $C3$ refer to the set of characteristics used to compute that variable (see section 3.1). Columns (1), (4) and (7) display the first stages of each 2SLS estimation. *Clinic* is a variable equal to 1 if the type of research is clinical, 0 otherwise. All regressions include dummies for Congress terms and type of research and their interactions. Cluster-robust standard errors are reported in parentheses. * p<0.05, + p<0.1.

Table 8: **Estimation of the effect of 1995 majority shift**

	All		*Clinical*	
Disease	**Actual**	**Counterfactual**	**Actual**	**Counterfactual**
Blood	5.41e+07	5.35e+07	1.86e+07	1.84e+07
CNS	4.81e+08	4.66e+08	1.96e+08	1.89e+08
Cardiovascular	3.26e+08	3.24e+08	1.40e+08	1.39e+08
Congenital	1.95e+08	2.20e+08	6.57e+07	7.60e+07
Digestive	1.96e+08	2.00e+08	7.77e+07	8.00e+07
Endocrine	4.71e+08	4.82e+08	1.25e+08	1.28e+08
HIV	8.04e+08	9.47e+08	3.43e+08	4.15e+08
Infectious	7.67e+08	8.07e+08	1.62e+08	1.72e+08
Psychiatric	5.08e+08	4.89e+08	3.44e+08	3.30e+08
Reproductive	3.24e+08	3.23e+08	1.51e+08	1.50e+08
Respiratory	8.77e+07	8.53e+07	3.53e+07	3.42e+07
Skin	1.90e+08	1.89e+08	9.29e+07	9.23e+07
Urinary	9.12e+07	9.35e+07	3.26e+07	3.35e+07
Total	3.46e+08	3.60e+08	1.37e+08	1.43e+08

Notes: the table presents the estimation of the actual and counterfactual average yearly amount of funds that would have been awarded through competing R grants between 1995 and 2002 for any type of research (*All*), and for clinical research only (*Clinical*), if the House majority had not changed in 1995. The counterfactual estimations rely on the specification of column (4) in Table 3.

Table 9: **Effect of state affiliation of congressmen on NIH funds**

	(1)	(2)	(3)	(4)	(5)
Dependent variable: Log of funds	by disease, by type of research and by state			by state	by type of research and by state
Number of state congressmen in:					
HouS	-0.040		-0.026	-0.018	-0.094
	(0.028)		(0.031)	(0.024)	(0.062)
HouS $\times Clinic$	-0.025		-0.023		0.14
	(0.047)		(0.052)		(0.10)
House Majority		-0.00015	-0.00091	-0.0043	-0.0086
		(0.0074)	(0.0075)	(0.0071)	(0.017)
House Maj. $\times Clinic$		0.0083	0.0063		0.015
		(0.012)	(0.013)		(0.026)
House Appropriations Com.			-0.020	-0.020	0.025
			(0.023)	(0.019)	(0.044)
House Appr. Com. $\times Clinic$			-0.0079		-0.088
			(0.036)		(0.068)
Senate Appropriations Subcom.			0.084	-0.029	0.022
			(0.063)	(0.052)	(0.070)
Senate Appr. Subcom. $\times Clinic$			0.014		-0.071
			(0.092)		(0.18)
Senate Majority			0.030	0.0052	0.060
			(0.019)	(0.025)	(0.055)
Senate Maj. $\times Clinic$			-0.017		-0.062
			(0.030)		(0.073)
Senate Appropriations Com.			-0.057	0.0020	-0.0012
			(0.047)	(0.039)	(0.076)
Senate Appr. Com. $\times Clinic$			0.081		0.054
			(0.070)		(0.13)
Senate Labor Com.			0.022	0.0073	0.098
			(0.033)	(0.032)	(0.071)
Senate Lab. Com. $\times Clinic$			-0.046		-0.21*
			(0.055)		(0.099)
Observations	10,212	10,212	10,212	400	796
R^2	0.017	0.017	0.018	0.105	0.198

Notes: this table reports the OLS estimation of the first difference of equation 5 for the period 1985-2002. The dependent variable *funds by disease, type of research and state* is the difference between the log of NIH funds for research of type θ on disease d awarded through R01 grants to research institutes in state s in Congress terms t and $t-1$. *Number of state congressmen in ...* is the difference between the number of congressmen from state s in ... in Congress terms t and $t-1$. Column (3) includes controls for the log of life-years lost in any group. Columns (4) and (5) aggregate data at the state level and state and type of research level respectively. All regressions include dummies for Congress terms and type of research and their interactions. Cluster-robust standard errors are in parentheses. * p<0.05, + p<0.1.

Table 10: **Effect of local life-years lost on research focus**

	(1)	(2)	(3)
Dependent variable is funds by disease, type of research and state			
Log of life-years lost in state s	0.095*	0.094*	0.055
	(0.046)	(0.046)	(0.057)
Log of life-years lost in state s $\times Clinic_\theta$			0.079
			(0.069)
Interaction Congress term-$Clinic_\theta$	No	Yes	Yes
Interaction Disease-$Clinic_\theta$	No	Yes	Yes
Observations	11,558	11,558	11,558
R^2	0.854	0.870	0.872

Notes: this table reports the OLS estimation of equation 6 for the period 1985-2002. The dependent variable *funds by disease, type of research and state* is the log of NIH funds for research of type θ on disease d awarded through R grants to research institutes in state s in Congress terms t and $t-1$. *Life-years lost in state s* is the log of life-years lost to 100 due to disease d in the districts/states represented in state s in Congress terms t. *Clinic* equals 1 if the type of research is clinical, 0 otherwise. All regressions include dummies for Congress terms, diseases, types of research and states. Cluster-robust standard errors in parentheses. * p<0.05, + p<0.1.

Table 11: **Life-years lost and *expected* NIH funds**

Characteristics	(1) $C1$	(2) $C1$	(3) $C1$	(4) $C2$	(5) $C2$	(6) $C2$	(7) $C3$	(8) $C3$	(9) $C3$
	Dependent variable is expected Log of funds by disease and type of research								
Log of life-years lost in:									
HouS	0.16+		-0.053	0.16+		-0.047	0.18*		0.0096
	(0.095)		(0.14)	(0.091)		(0.13)	(0.079)		(0.090)
HouS $\times Clinic$			0.36+			0.34+			0.27+
			(0.20)			(0.19)			(0.14)
House Majority		0.65*	0.74*		1.39*	1.60*		0.25+	0.22*
		(0.25)	(0.30)		(0.59)	(0.57)		(0.12)	(0.089)
House Maj. $\times Clinic$			-0.38			-0.60			-0.075
			(0.60)			(1.30)			(0.22)
Observations	208	208	208	208	208	208	208	208	208
R-squared	0.871	0.872	0.873	0.871	0.872	0.873	0.872	0.872	0.873

Notes: this table reports the OLS estimation of the first difference of equation 4 for the period 1985-2002, where the dependent variable is the *expected* log of funds by disease as defined in section 6.3, separately for each estimation of the life-years lost variable at the district level. $C1$, $C2$ and $C3$ refer to the set of characteristics used to compute that variable (see section 3.1). *Clinic* is a variable equal to 1 if the type of research is clinical, 0 otherwise. All regressions include dummies for Congress terms and type of research and their interactions. Cluster-robust standard errors are reported in parentheses. * $p<0.05$, + $p<0.1$.

9 Appendix

This section presents a simple static model that aims to formalize the description of the NIH budget process. Let K denote the set of "research categories" partitioning the NIH budget. For instance, K could be the set of diseases, the set of research institutes, the combination of both, or any other set relevant for the classification of research projects. For any research category $k \in K$, the level of public investment $F_k \in \mathbb{R}^+$. Investing F_k in research k yields a scientific return equal to $a_k F_k^\gamma$, where $\gamma \in (0,1)$ and the productivity factor a_k is independently drawn from a random variable with density $g_k(a_k)$, with a support in $\mathbb{R}^+$.

Agents and Information

Let C be the set of legislators deciding on the distribution of funds $F_k \in \mathbb{R}^+$, and N, an additional agent that is not part of C. The set C may be part or the whole Congress, and may include the President. I assume that the members of C do not know a_k, whereas N does. In addition, I assume that C has the possibility to delegate the allocation of a chosen amount of funds into the categories of a chosen subset M of K to the NIH N, and decides of the amount of funds for the other categories $L \equiv K \setminus M$.

Preferences

The utility an agent $i \in C \cup \{N\}$ derives from the allocation $((F_k)_{k\in K}$ is:

$$U^i\Big((F_k)_{k\in K}\Big) \equiv \sum_{k\in K} a_k F_k^\gamma A_k^i - \lambda \sum_{k\in K} F_k \tag{8}$$

where $\forall k \in K, A_k^i > 0, \lambda > 0$. I assume that, at the end of the budget process, the allocation of funds maximizes a weighted average utility of the members of C:

$$\sum_{i\in C} w^i U^i\Big((F_k)_{k\in K}\Big) \tag{9}$$

with $\sum_{i\in C} w^i = 1$. The allocation of funds across research categories thus solves the following problem:

$$\begin{cases} Max_{\{F\in\mathbb{R}^+, L\subset K, \mathbf{F}_L\in\mathbb{R}^{+\ \#L}\}}\mathbf{E}\left[\sum_{i\in C}\sum_{k\in L} a_k F_k^{\gamma} B_k^i + \sum_{k\in M} a_k \hat{F}_k^{\gamma} B_k^i \middle| \mathbf{g}\right] - \lambda F \\ s.t. \begin{cases} \hat{\mathbf{F}}_M \equiv Argmax_{\{\mathbf{F}_M\in\mathbb{R}+^{\#M}\}} \sum_{k\in M} a_k F_k^{\gamma} {B_k}^N \\ \sum_{k\in M} F_k = F - \sum_{k\in L} F_k \end{cases} \\ \sum_{k\in L} F_k + \sum_{k\in M} \hat{F}_k = F \\ L\cup M = K, L\cap M = \emptyset \end{cases} \tag{10}$$

where: $\mathbf{F}_L \equiv (F_k)_{k\in L}$ for any $L \subset K$ and $\mathbf{g} \equiv (g_k)_{k\in K}$. A solution to that problem always exists, since a solution of the sub-problem defined by adding the constraint $L = L_0$ to the problem 10 has a solution for any $L_0 \subseteq K$, and there is a finite number of such subsets. The solution need not be unique, however. Let $\left(F^*, L^*, \mathbf{F}^*_{L^*}\right) \in \mathbb{R}^+ \times 2^K \times \mathbb{R}^{+\ \#L^*}$ denote a solution of the previous program.

Proposition 1. *Let $K \equiv \{1, ..., n\}$ the set of research categories, $g_k(a)$, the probability density function of the productivity factor a_k, and assume that g_2 second-order dominates g_1 and $B_1^i = B_2^i$, for $i \in C \cup \{N\}$.*

If $1 \in L^*$ for some solution $\left(F^*, L^*, \boldsymbol{F}^*_{L^*}\right) \in \mathbb{R}^+ \times 2^K \times \mathbb{R}^{+\ \#L^*}$ of the Congress' problem, then there also exists a solution $\left(\tilde{F}^*, \tilde{L}^*, \boldsymbol{F}^*_{\tilde{L}^*}\right) \in \mathbb{R}^+ \times 2^K \times \mathbb{R}^{+\ \#\tilde{L}^*}$ of this problem such that $2 \in \tilde{L}^*$.

Proof. Consider the solution $\left(L^*, \mathbf{F}^*_{L^*}\right)$ defined in the proposition. If $2 \in L^*$, then taking $\left(\tilde{L}^*, \mathbf{F}^*_{\tilde{L}^*}\right) = \left(L^*, \mathbf{F}^*_{L^*}\right)$ proves the claim. If $2 \notin L^*$, let the function $V : [0,1] \to \mathbb{R}$ be defined as:

$$\begin{aligned} V(\delta) \quad = \quad & \int_0^{+\infty} \dots \int_0^{+\infty} \left((1-\delta) c_\alpha A_1^C {F_1^*}^{\gamma} + c_\alpha \delta A_1^C \hat{F}_\alpha^{\gamma}\right) g_1(c_\alpha) dc_\alpha g_2(c_2) dc_2 g_3(c_3) dc_3 \dots g_n(c_n) dc_n + \\ & \int_0^{+\infty} \dots \int_0^{+\infty} \left(\delta c_\alpha A_2^C {F_1^*}^{\gamma} + (1-\delta) c_\alpha A_2^C \hat{F}_\alpha^{\gamma}\right) g_1(c_1) dc_1 g_2(c\alpha) dc\alpha g_3(c_3) dc_3 \dots g_n(c_n) dc_n + \\ & \sum_{k\in L^*\setminus\{1\}} \int_0^{+\infty} a_k {F_k^*}^{\gamma} A_k^C g_k(a_k) da_k + \\ & \sum_{k\in K\setminus L^*\cup\{2\}} \int_0^{+\infty} \dots \int_0^{+\infty} a_k \hat{F}_k^{\gamma} A_k^C g_1(c_1) dc_1 g_2(c_2) dc_2 \dots g_n(c_n) dc_n \end{aligned}$$

where $\hat{F}_\alpha, \hat{\mathbf{F}}_{K\setminus(L^*\cup\{2\})}$ solves:

$$\begin{cases} Max_{\{F_\alpha\in\mathbb{R}^+, F_k\in\mathbb{R}^+, k\in K\setminus(L^*\cup\{2\})\}} c_\alpha F_\alpha^{\gamma} A_1^N + \sum_{k\in K\setminus(L^*\cup\{2\})} a_k F_k^{\gamma} A_k^N \\ s.t.\ F_\alpha + \sum_{K\setminus(L^*\cup\{2\})} F_k = F^* - \sum_{k\in L^*} F_k^* \end{cases} \tag{11}$$

$V(0)$ is C's utility achieved at the solution $\left(F^*, L^*, \mathbf{F}^*_{L^*}\right)$, and $V(1)$ is the utility that would be achieved if 1 and 2 were swapped, every other variable chosen by C remaining unchanged (note that the expected allocation of funds across the research categories of $K \setminus (L^* \cup \{2\})$ by N is then the same before and after the swap). The first-order conditions of N's problem impose, for any $k \in \{\alpha\} \cup K \setminus (L^* \cup \{2\})$:

$$\hat{F}_k = \frac{(a_k B_k)^{\frac{1}{1-\gamma}}}{\sum_{k \in \{c\} \cup K \setminus (L^* \cup \{2\})} (a_k B_k)^{\frac{1}{1-\gamma}}} \left(F^* - \sum_{k \in L^*} F^*_k\right)$$

so that:

$$V'(\delta) = A_1^C \int_0^{+\infty} \dots \int_0^{+\infty} c_\alpha \hat{F}_c^\gamma [g_1(c_\alpha) - g_2(c_\alpha)] dc_\alpha \Pi_{k \in K \setminus (L^* \cup \{2\})} g_k(a_k) da_k$$

is nonnegative since the function $c_\alpha \mapsto c_\alpha \hat{F}_\alpha^\gamma$ is convex and g_2 second-order dominates g_1. V is nondecreasing in δ, $V(0) \leq V(1)$, so that the solution defined by the swap of 1 and 2 is indeed an optimum of C's maximization problem, in which 2 is not delegated.

□

The proposition states that C is more likely to delegate the allocation of funds across research categories about which it has less precise information. The intuition for this result is similar to Aghion and Tirole 1997's distinction of forlam versus real authority.

Clinical research is more applied than basic research, and its potential outcomes are bounded by the law due to its use of human data. Congressmen should thus have a more precise information about clinical than basic research productivity. Prop 1 in this context thus implies that they will be more likely to delegate the allocation decision of basic research funds to the NIH.

Table 12: **National Institutes of Health in 2002**

National Cancer Institute
National Heart Lung and Blood Institute
National Institute of Dental and Craniofacial Research
National Institute of Diabetes and Digestive and Kidney Diseases
National Institute of Neurological Disorders and Stroke
National Institute of Allergy and Infectious Diseases
National Institute of General Medical Sciences
Eunice Kennedy Shriver National Institute of Child Health and Human Development
National Eye Institute
National Institute of Environmental Health Sciences
National Institute on Aging
National Institute of Arthritis and Musculoskeletal and Skin Diseases
National Institute on Deafness and Other Communication Disorders
National Institute of Mental Health
National Institute on Drug Abuse
National Institute on Alcohol Abuse and Alcoholism
National Institute of Nursing Research
National Human Genome Research Institute
National Institute of Biomedical Imaging and Bioengineering
National Center for Research Resources
National Center for Complementary and Alternative Medicine
National Institute on Minority Health and Health Disparities
John E. Fogarty International Center for Advanced Study in the Health Sciences
National Library of Medicine
Office of the Director
Building and Facilities

Table 13: **Summary statistics by disease**

Diseases	#Deaths (/1000)	Log of life-years lost			Log of funds	
		House	House Majority	HouS	Basic	Clinical
Blood	33.9	7.8	7.79	7.8	18.5	17.7
			.0403	.00987	.341	.454
CNS	48	8.05	8.02	8.06	20.5	20
			.0169	.00543	.354	.474
Cardiovascular	754	10.6	10.6	10.6	20.1	19.8
			.0555	.0272	.362	.355
Congenital	25	8.59	8.49	8.59	19.9	19.1
			.0691	.0855	.221	.262
Digestive	158	9.29	9.31	9.28	19.8	19
			.0339	.0382	.247	.53
Endocrine	52.2	8.14	8.12	8.14	20.9	19.7
			.0287	.0457	.225	.333
HIV	18.8	7.88	7.77	7.85	21.1	20.6
			.207	.129	.222	.414
Infectious	96.6	8.59	8.56	8.59	21.3	19.9
			.0796	.0644	.277	.381
Psychiatric	40.5	8.38	8.28	8.38	19.9	20.6
			.0547	.00936	.406	.472
Reproductive	80.6	8.62	8.63	8.62	20	19.7
			.0319	.0172	.382	.547
Respiratory	213	9.57	9.57	9.58	18.9	18.2
			.0471	.0123	.31	.535
Skin	20.8	7.34	7.31	7.34	19.5	19.4
			.0381	.015	.281	.388
Urinary	32.7	7.58	7.59	7.58	19.1	18.3
			.0403	.0446	.226	.417

Notes: columns 1 to 4 report the total number of deaths (in thousands), and normalized log of life-years lost in the US in 1989, and the average normalized log of life-years lost in the House Appropriations Subcommittee for the NIH, and in the House Majority for the whole period over the period 1985-2002, as defined in equation 3. *CNS* comprises any nervous system disorder, *Infectious* includes any infection that is not sexually transmitted, *HIV* comprises sexually-transmitted infections, *Endocrine* comprises diabetes as well as other endocrine and metabolic disorders, *Blood/Lymphatic* comprises blood/lymphatic cancers *only*, *Psychiatric* comprises mainly self-destructive behavioral disorders, *Congenital* comprises congenital and developmental disorders. The funds are the NIH funds awarded through R grants in the period 1985-2002.