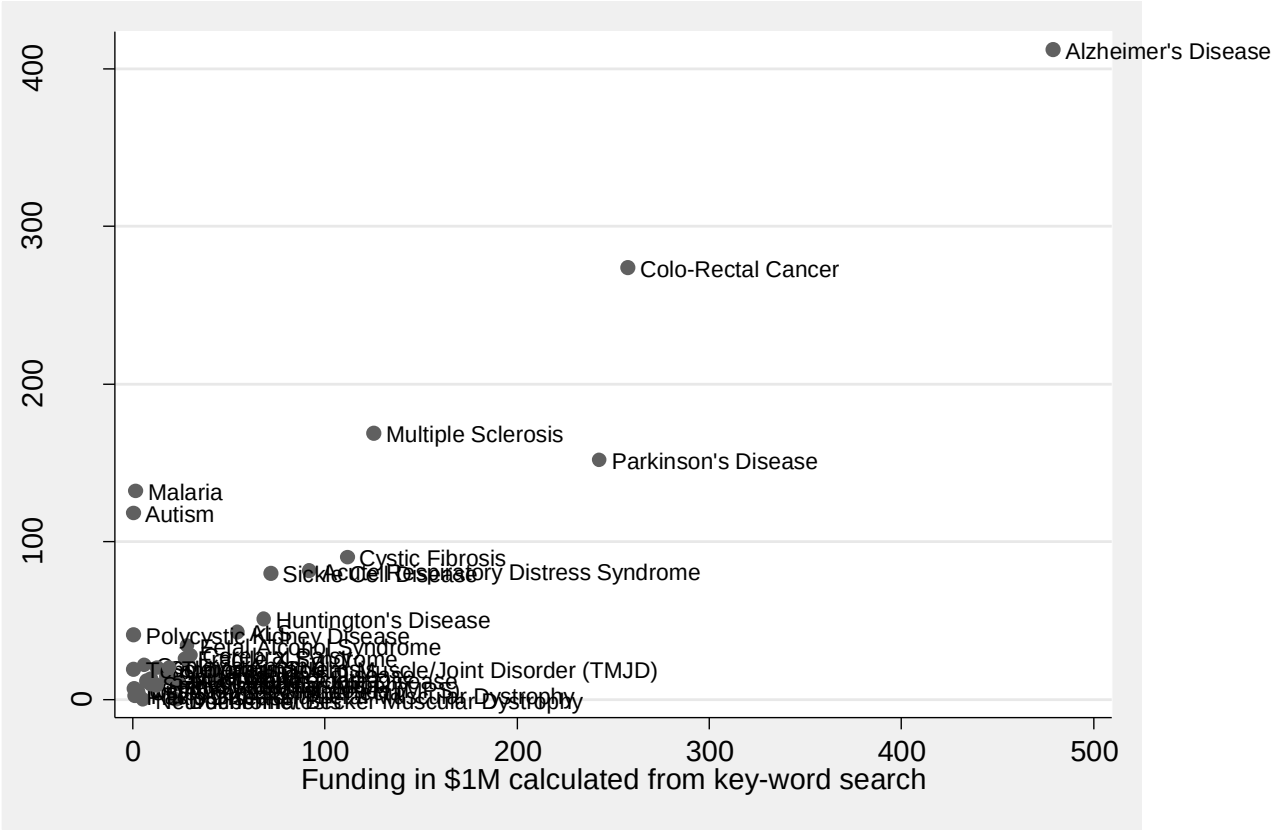
Online Appendix

Can private money buy public science? Disease group lobbying and federal funding for biomedical research

Deepak Hegde[†] **and Bhaven Sampat**[‡]

Figure A1: Keyword search algorithm v/s NIH-RCDC

Figure shows NIH funding numbers for 33 rare diseases as calculated by our keyword search (on the X-axis) and as reported by NIH's Research and Conditions Disease Categorization (RCDC) initiative for the year 2008. Both funding numbers are in millions of dollars.



[†] New York University; email: dhegde@stern.nyu.edu

[‡] Columbia University; email: bns3@columbia.edu

Figure A2: Year-to-year within disease changes in lobbying expenditures

Figure presents box plots of the distribution of year-to-year changes in within-disease lobbying expenditures calculated as the percentage difference between previous year and current year lobbying expenditures associated with each disease. The upper and lower hinges of each box represents the 25th and 75th percentiles of the percentage differences for each year and the line in the box represents the median difference. The upper and lower whiskers depict the upper and lower adjacent values for the years respectively. Years with more "spread out" distributions (both below and above zero) are more likely to be those in which there was a larger change in the cross-disease composition of lobbying expenditures.

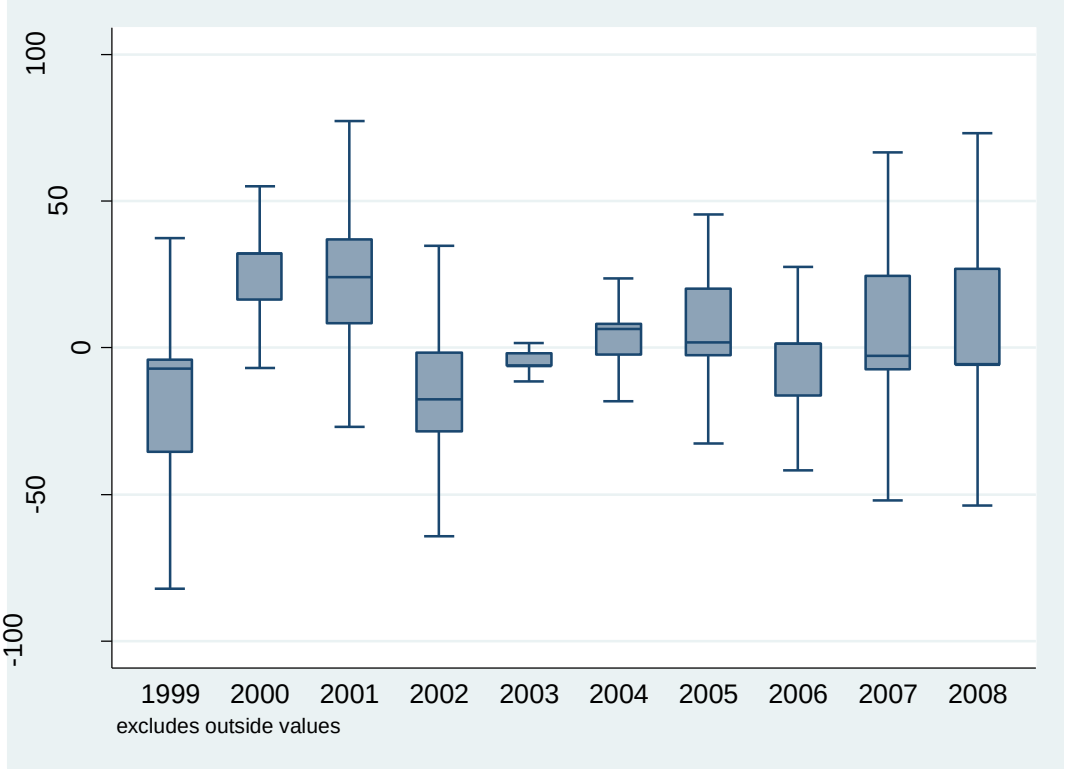


Table A1: Research-related lobbying interest groups

The following list of lobbying organizations was compiled from the disclosures of lobbying activities filed with the Secretary of the Senate's Office of Public Records (SOPR) and available from the Center for Responsive Politics. The organizations had spent at least $10,000 in one of the six-month periods on lobbying between 1998 and 2008, and disclosed some interest in research on at least one of the 955 rare diseases in our sample.

ALZHEIMER'S ASSN	INTERSTITIAL CYSTITIS ASSN
AMERICAN CANCER SOCIETY	IRON DISORDERS INSTITUTE
AMERICAN FOUNDATION FOR AIDS RESEARCH	JEFFREY MODELL FOUNDATION
AMERICAN HEART ASSN	KENNEDY KRIEGER INSTITUTE
AMERICAN LIVER FOUNDATION	KIDNEY CANCER ASSN
AMERICAN SOCIAL HEALTH ASSN	LANCE ARMSTRONG FOUNDATION
AMERICAN SPINAL INJURY ASSN	LYMPHOMA RESEARCH FOUNDATION
AMERICAN TINNITUS ASSN	MULTIPLE MYELOMA RESEARCH FOUNDATION
ARTHRITIS FOUNDATION	MUSCULAR DYSTROPHY ASSN
AUTISM SOCIETY OF AMERICA	MYOSITIS ASSN
AUTISM SPEAKS	NATIONAL ALLIANCE FOR THE MENTALLY ILL
CHARCOT-MARIE-TOOTH ASSN	NATIONAL DOWN SYNDROME CONGRESS
CHILDREN'S TUMOR FOUNDATION	NATIONAL FRAGILE X FOUNDATION
CHRISTOPHER REEVE PARALYSIS FOUNDATION	NATIONAL KIDNEY FOUNDATION
CROHN'S & COLITIS FOUNDATION OF AMERICA	NATIONAL MARFRAN FOUNDATION
CYSTIC FIBROSIS FOUNDATION	NATIONAL MULTIPLE SCLEROSIS SOCIETY
DEAFNESS RESEARCH FOUNDATION	NATIONAL SLEEP FOUNDATION
DYSTONIA MEDICAL RESEARCH FOUNDATION	NATIONAL STROKE ASSN
EYE BANK ASSN OF AMERICA	NATL CENTER FOR LEARNING DISABILITIES
FRIEDREICHS ATAXIA RESEARCH ALLIANCE	NEUROFIBROMATOSIS FOUNDATION
FRIENDS OF CANCER RESEARCH	NEUROPATHY ASSN
GENETIC ALLIANCE	SCHEPENS EYE RESEARCH INSTITUTE
HUNTINGTON'S DISEASE SOCIETY OF AMERICA	SCLERODERMA FOUNDATION
IMMUNE DEFICIENCY FOUNDATION	SPINA BIFIDA ASSN OF AMERICA
INTERNATIONAL MYELOMA FOUNDATION	TRIGEMINAL NEURALGIA ASSN
INTERNATIONAL RETT SYNDROME ASSN	TUBEROUS SCLEROSIS ALLIANCE

Table A2: Number of publications and scientific opportunity

Table reports the annual number of publications associated with each disease, and the years in which scientific advances associated with the disease were reported in media articles. We also reference the earliest article (between 1998 and 2006) that reported the advance corresponding to the disease below.

Disease	1998	1999	2000	2001	2002	2003	2004	2005	2006	Years of advance
ALEXANDER DISEASE (a)	0	0	0	0	5	15	9	9	18	2001, 06
ATRIAL SEPTAL DEFECTS (b)	207	210	206	268	258	308	347	323	383	2002, 06
COLITIS, COLLAGENOUS (c)	0	0	0	0	0	0	5	20	24	2003, 06
DISTAL MYOPATHY (d)	0	0	0	0	0	0	4	12	10	2004, 05
ENCEPHALITIS, HERPES SIMPLEX (e)	2	17	37	55	51	58	54	58	54	2003,04,05
ERYTHROPOIETIC PROTOPORPHYRIA (f)	2	5	4	2	2	0	3	19	20	2005
FACIOSCAPULOHUMERAL MUSCULAR DYSTROPHY (g)	0	12	11	16	23	22	31	29	26	2002,04
GASTROINTESTINAL STROMAL TUMORS (h)	0	0	1	0	0	6	89	277	311	2005,06
RETINOSCHISIS (i)	0	0	0	1	7	21	19	28	29	2002,05
RETT SYNDROME (j)	50	52	54	122	72	82	51	101	94	1999
HERMANSKY PUDLAK SYNDROME (k)	0	1	6	17	20	21	22	25	22	2005
NEUROMYOTONIA (l)	0	6	18	7	8	3	9	9	24	2003, 06

(a) "Genetic Basis of Disease Discovered," *Pain & Central Nervous System Week*, 27 January 2001

(b) "Making repairs without lifting the hood," *The Globe and Mail*, 19 November 2002

(c) "Collagenous colitis with pseudomembrane formation may occur with Ig deficiency," *Health and Medicine Week*, 10 November 2003

(d) "Mutations toward the 3' end of MYH7 cause early-onset distal myopathy" Obesity, *Fitness & Wellness Week*, 13 November 13 2004

(e) "HSV Encephalitis; Pediatric encephalitis relapses may be due to HSV replication or inflammation" Obesity, *Drug Week*, 26 September 2003

(f) "ABCG2 has role in regulating PPIX levels during erythroid differentiation" *Life Science Weekly*, 24 May 2005

(g) "Muscular Dystrophy: Scientists identify a new kind of genetic problem," *Genomics & Genetics Weekly* 27 September 2002

(h) "Gastrointestinal Stromal Tumors; Radiologic findings of ruptured gastrointestinal stromal tumors described" *Cancer Weekly*, 27 Dec 2005

(i) "Vitreous detachments and retinal holes," *Optician*, 27 Sep 2002

(j) "Rett syndrome is caused by mutations in X-linked MECP2," http://www.rsrt.org/rett-and-mecp2-disorders/rett-syndrome/

(k) "Lung pathology studied, correlation with interstitial inflammation revealed," *Science Letter*, 3 MAY 2005

(l) "Neuroimmunology; Myokymia, dermatomyositis cause may be antibodies to voltage-gated K+ channels*" Medical Devices & Surgical Technology Week*, 24 AUG 2003

Table A3: Relationship between soft earmarks and lobbying

Table reports estimates of the effect of previous year lobbying expenditures of research-related interest groups, NIH grants, deaths, and publications on earmarks. Column 1 reports OLS estimates obtained by specifying the dependent variable as the number of earmarks. Column 2 reports OLS estimates obtained by specifying the dependent variable as log(1+ the number of earmarks). Column 3 reports Logit estimates obtained by specifying the dependent variable as a dummy variable that indicates the presence of at least an earmark (for the disease-year). Column 4 reports Poisson pseudo-maximum-likelihood (PPML) estimates obtained by specifying the dependent variable as the number of earmarks. The unit of observation for all estimations is disease-year, and the sample has observations for 955 diseases for each year between 1998 and 2008. All S.E.s, in brackets, are clustered at the disease-level: + p<0.1; * p<0.05; ** p<0.01.

Column #	1	2	3	4
Specification	OLS	OLS	Logit	PPML
D.V.	N. of earmarks	log(1+N of earmarks)	Earmarks (0/1)	N. of earmarks
Log Lobbying expenditures (1-year lagged)	0.246*	0.059*	0.581*	0.388**
	[0.115]	[0.027]	[0.288]	[0.091]
Log Deaths (1-year lagged)	0.008	-0.003	-0.416	-0.332*
	[0.019]	[0.006]	[0.331]	[0.148]
Log Publications (1-year lagged)	0.027	0.013	1.079+	0.488+
	[0.022]	[0.010]	[0.596]	[0.274]
Log NIH grants (1-year lagged)	0.001	0.001	0.021	0.03
	[0.002]	[0.001]	[0.111]	[0.050]
Constant	-0.14	-0.02	-6.21	-3.26
Disease fixed-effects (955)	Y	Y	Y	Y
Year fixed-effects (10)	Y	Y	Y	Y
Observations	9,550	9,550	9,550	9,550
Adjusted R-square	0.606	0.660		
Log Pseudo Likelihood			-355.4	-842.0

Table A4: Yearly estimates of the relationship between earmarks and lobbying

Table reports ordinary least squares estimates of the effects of lobbying on congressional soft earmarks for each of the years during the 1999-2008 period. The unit of observation in the estimation sample is disease-year, and the sample has observations for 955 diseases for each year between 1998 and 2008. All S.E.s, in brackets, are clustered at the disease-level; + p<0.1; * p<0.05; ** p<0.01.

D.V.= No. of earmarks	
Log Lobbying Exp. X Y1999	0.230+
	[0.123]
Log Lobbying Exp. X Y2000	0.273+
	[0.140]
Log Lobbying Exp. X Y2001	0.285*
	[0.133]
Log Lobbying Exp. X Y2002	0.220+
	[0.115]
Log Lobbying Exp. X Y2003	0.206+
	[0.112]
Log Lobbying Exp. X Y2004	0.263*
	[0.125]
Log Lobbying Exp. X Y2005	0.252*
	[0.120]
Log Lobbying Exp. X Y2006	0.272*
	[0.126]
Log Lobbying Exp. X Y2007	0.260*
	[0.118]
Log Lobbying Exp. X Y2008	0.244*
	[0.115]
Log Deaths (1-year lagged)	0.007
	[0.019]
Log Publications (1-year lagged)	0.026
	[0.022]
Log NIH grants (1-year lagged)	0.001
	[0.002]
Constant	-0.128
Disease fixed-effects (955)	Y
Year fixed-effects (10)	Y
Observations	9,550
Adjusted R-square	0.608

Table A5: Relationship between earmarks and lobbying by Congressional chamber

Table reports Ordinary Least Squares (OLS) estimates of the effect of lobbying by research-related interest groups on the number of current year earmarks. Column 1 reports estimates of the effects of the explanatory variables on the earmarks by both chambers, Column 2 on earmarks by the House appropriations committee, and Column 3 on earmarks by the Senate appropriations committee. The unit of observation in the estimation sample is disease-year, and the sample has observations for 955 diseases for each year between 1998 and 2008. All S.E.s, in brackets, are clustered at the disease-level; + p<0.1; * p<0.05; ** p<0.01.

Column #	1	2	3
D.V.= No. of earmarks	Both Chambers	House	Senate
Log Lobbying expenditures (1-year lagged)	0.246*	0.096*	0.150*
	[0.115]	[0.049]	[0.072]
Log Deaths (1-year lagged)	0.008	-0.008	0.016
	[0.019]	[0.007]	[0.016]
Log Publications (1-year lagged)	0.027	0.007	0.02
	[0.022]	[0.006]	[0.017]
Log NIH grants (1-year lagged)	0.001	0.001	0
	[0.002]	[0.001]	[0.001]
Constant	-0.14	-0.03	-0.11
Disease fixed-effects (955)	Y	Y	Y
Year fixed-effects (10)	Y	Y	Y
Observations	9,550	9,550	9,550
Adjusted R-square	0.606	0.482	0.606

Table A6: Size and duration of unsolicited projects and solicited projects

Table shows the average size (in millions of 2010 dollars), duration in years, and proportion of grants that were awarded for "centers" (rather than research projects) under each of the two distinct NIH grant mechanisms (Request for applications (RFAs) and program announcements (PA) grant mechanisms, and grants for unsolicited investigator-initiated projects (II). Duration is calculated as the total number of years for which each grant is awarded (through continuation of the original award). Since we do not have data on grants after 2008, both average grant size and durations are subject to censoring for the later years.

Year	Average grant size		Average grant duration		Share of "center" grants	
	II-grants	RF+PA grants	II-grants	RF+PA grants	II-grants	RF+PA grants
1998	1,094,517	3,210,097	3.7	4.4	2.3%	16.9%
1999	1,333,978	3,665,024	4.1	5.0	2.8%	12.6%
2000	1,161,284	2,615,332	3.8	4.3	1.3%	5.7%
2001	1,164,530	2,260,154	3.6	4.3	1.2%	8.2%
2002	1,283,886	1,446,020	3.6	3.8	2.0%	1.7%
2003	1,202,673	3,227,665	3.5	3.8	2.2%	5.4%
2004	1,117,947	1,974,653	3.3	3.2	1.3%	6.5%
2005	949,604	2,851,004	2.9	3.3	1.4%	4.9%
2006	741,512	1,168,044	2.4	2.6	1.4%	4.0%
2007	547,224	964,521	1.8	1.9	0.9%	6.4%
2008	295,937	470,329	1.0	1.0	0.7%	1.9%
1998-2008	961,057	2,088,348	3.0	3.3	1.8%	5.9%

Table A7: The relationship between lobbying and disease-characteristics (with alternative measures of disease burden and scientific opportunity)

Table reports Ordinary Least Squares (OLS) estimates of the effect of lagged deaths, hospital utilization, publications (weighted by the impact factors of the corresponding journals) and NIH funding on current year expenditures of research-related lobbying interest groups. The unit of observation in the estimation sample is disease-year, and the sample has observations for 955 diseases for each year between 1998 and 2008. All S.E.s, in brackets, are clustered at the disease-level: + p<0.1; * p<0.05; ** p<0.01.

Column #	1	2	3	4	5
D.V.= Log Lobbying expenditures					
Log Deaths (1-year lagged)	0.066**	0.052*	0.051*	0.051*	0.051*
	[0.021]	[0.024]	[0.024]	[0.024]	[0.024]
Log Hospital Utilization (1-year lagged)	0.004	-0.003	-0.003	-0.011	
	[0.009]	[0.010]	[0.010]	[0.009]	
Log JIF-weighted Pubs. (1-year lagged)	0.052**	0.021*	0.019*	0.015+	
	[0.019]	[0.009]	[0.009]	[0.009]	
Log Pubs. (5-year Stock)					0.013+
					[0.007]
Log NIH grants (1-year lagged)			0.009**	0.005	0.005+
			[0.003]	[0.003]	[0.003]
Constant	0.374	0.505	0.482	0.412	0.364
Disease fixed-effects (955)	N	Y	Y	Y	Y
Year fixed-effects (10)	N	N	N	Y	Y
Observations	9,550	9,550	9,550	9,550	9,550
Adjusted R-square	0.038	0.904	0.905	0.907	0.907

Table A8: The relationship between earmarks and lobbying (with alternative measures of disease burden and scientific opportunity)

Table reports OLS estimates of the effect of lagged deaths, hospital utilization, and publications (weighted by the impact factors of the corresponding journals) and NIH funding on the number of current year earmarks. The unit of observation in the estimation sample is disease-year, and the sample has observations for 955 diseases for each year between 1998 and 2008. All S.E.s, in brackets, are clustered at the disease-level: + p<0.1; * p<0.05; ** p<0.01.

Column #	1	2	3	4	5	6
Specification	OLS	OLS	OLS	OLS	OLS	OLS
D.V.= No. of earmarks						
Log Lobbying expenditures (1-year lagged)		0.043+	0.244*	0.244*	0.247*	0.246*
		[0.023]	[0.114]	[0.114]	[0.115]	[0.115]
Log Deaths (1-year lagged)	0.039**	0.036**	0.006	0.006	0.008	0.009
	[0.013]	[0.013]	[0.019]	[0.019]	[0.019]	[0.019]
Log Hospital Utilization (1-year lagged)	0.003	0.003	0.018+	0.018+	0.016	
	[0.004]	[0.004]	[0.010]	[0.010]	[0.010]	
Log JIF-weighted Pubs. (1-year lagged)	0.034**	0.032**	0.015	0.015	0.014	
	[0.008]	[0.008]	[0.013]	[0.013]	[0.013]	
Log Pubs. (5-year Stock)						0.011
						[0.008]
Log NIH grants (1-year lagged)				0.001	0.001	0.001
				[0.002]	[0.002]	[0.002]
Constant	-0.072	-0.088	-0.182	-0.185	-0.182	-0.118
Disease fixed-effects (955)	N	N	Y	Y	Y	Y
Year fixed-effects (10)	N	N	N	N	Y	Y
Observations	9,550	9,550	9,550	9,550	9,550	9,550
Adjusted R-square	0.05	0.056	0.605	0.605	0.607	0.606

Table A9: The relationship between NIH grants and earmarks (with alternate measures of disease burden and scientific opportunity)

Table presents ordinary least squares (OLS) estimates of the effect of past year (logged) earmarks and (logged) lobbying expenditures on (logged) current year NIH funding for diseases. Columns 1-8 provide estimates of the effects of the explanatory variables on all new NIH funding. RFAs are grants through requests for applications and PAs are grants through program announcements. II indicates grants for unsolicited investigator-initiated projects. The unit of observation in the estimation sample is disease-year, and the sample has observations for 955 diseases for each year between 1998 and 2008. All S.E.s, in brackets, are clustered at the disease-level: + p<0.1; * p<0.05; ** p<0.01.

Column #	1	2	3	4	5	6	7	8
D.V.= Log NIH grants	All	RF+PA	PA	II	All	RF+PA	PA	II
Earmarks (1-year lagged)	0.054+	0.165**	0.214**	0.047	0.056+	0.167**	0.216**	0.048
	[0.031]	[0.060]	[0.061]	[0.043]	[0.031]	[0.060]	[0.061]	[0.043]
Log Lobbying expenditures (1-year lagged)	-0.004	0.001	0.059	0.004	-0.004	0.001	0.056	0.003
	[0.053]	[0.062]	[0.067]	[0.058]	[0.053]	[0.062]	[0.067]	[0.058]
Log Deaths (1-year lagged)	-0.037	0.112	0.084	-0.029	-0.036	0.113	0.085	-0.028
	[0.077]	[0.099]	[0.074]	[0.072]	[0.077]	[0.099]	[0.073]	[0.072]
Log Hospital Utilization (1-year lagged)	0.002	0.002	0.034+	0.024				
	[0.026]	[0.025]	[0.019]	[0.025]				
Log JIF-weighted Pubs. (1-year lagged)	0.077+	0.070*	0.037	0.044				
	[0.042]	[0.034]	[0.027]	[0.043]				
Log Pubs. (5-year Stock)					0.004	0.008	-0.028+	-0.018
					[0.019]	[0.018]	[0.016]	[0.018]
Log NIH grants (1-year lagged)	-0.047**	-0.009	0	-0.035*	-0.046**	-0.008	0.001	-0.034*
	[0.015]	[0.009]	[0.006]	[0.014]	[0.015]	[0.009]	[0.006]	[0.014]
Constant	1.799	0.374	0.194	1.582	1.981	0.536	0.423	1.785
Disease fixed-effects (955)	N	Y	Y	Y	Y	Y	Y	Y
Year fixed-effects (10)	Y	Y	Y	Y	Y	Y	Y	Y
Observations	9,550	9,550	9,550	9,550	9,550	9,550	9,550	9,550
Adjusted R-square	0.693	0.591	0.572	0.674	0.693	0.591	0.572	0.674