APPENDIX MATERIALS

Supporting Descriptives

This section describes supplemental descriptives and analysis relevant to my proximity and quality measures.

Appendix Figure A plots the full distribution of previous citations and publications for permanent and temporary reviewers. Specifically, past citations are defined as the number of citations, to 2008, for publications published by the reviewer in the 5 years prior to the grant review meeting. Past publications simply count the number of such publications. I find that, overall, permanent and temporary reviewers have similar qualifications although the two distributions are statistically different. See Table 3 in the main text for additional details about these distributions.

Appendix Figure B plots this distribution separately by funded and unfunded candidates. Although funded applicants have higher quality in terms of both publications and citations, this figure clearly shows that there are still many unfunded applications that go on to generate many publications and citations.

Appendix Table A provides additional descriptives comparing the distribution of measured application quality for funded and unfunded applicants. The top panel compares citation-based quality measures for funded and unfunded applicants at the mean, and 1st, 5th, 10th, 25th, 50th, 75th, 90th, 95th, and 99th percentiles. Funded applicants have consistently higher measures of quality than unfunded applicants, although it is important to note that many unfunded applications have high ex ante quality as captured by this text-matching approach. The bottom panel of Table A displays the same comparison for publications. Here, there are fewer differences between funded and unfunded applicants. The mean number of publications is higher for funded applicants, and this is driven by publications at the tail.

Appendix Figure C plots distributions of application quality by proximity to permanent reviewers. The top two panels plot the distributions of citations and publications for applicants cited by exactly one reviewer. For each graph, the solid line shows the distribution of quality among applicants cited by one permanent reviewer and the dotted line does so for those cited by one temporary reviewer. These distributions are statistically indistinguishable: a Kolmogorov–Smirnov test cannot reject the null that these two distributions are equal. Similarly, the upper-right-hand panel shows the same, but with quality measured using the number of publications associated with a grant. The bottom two panels of Appendix Figure C repeat this exercise for applicants who have been cited by a total of two reviewers. In this case, there are now three possibilities: the applicant has been cited by two temporary reviewers, two permanent, or one of each. In all of these cases, the distribution of applicant quality is statistically similar.

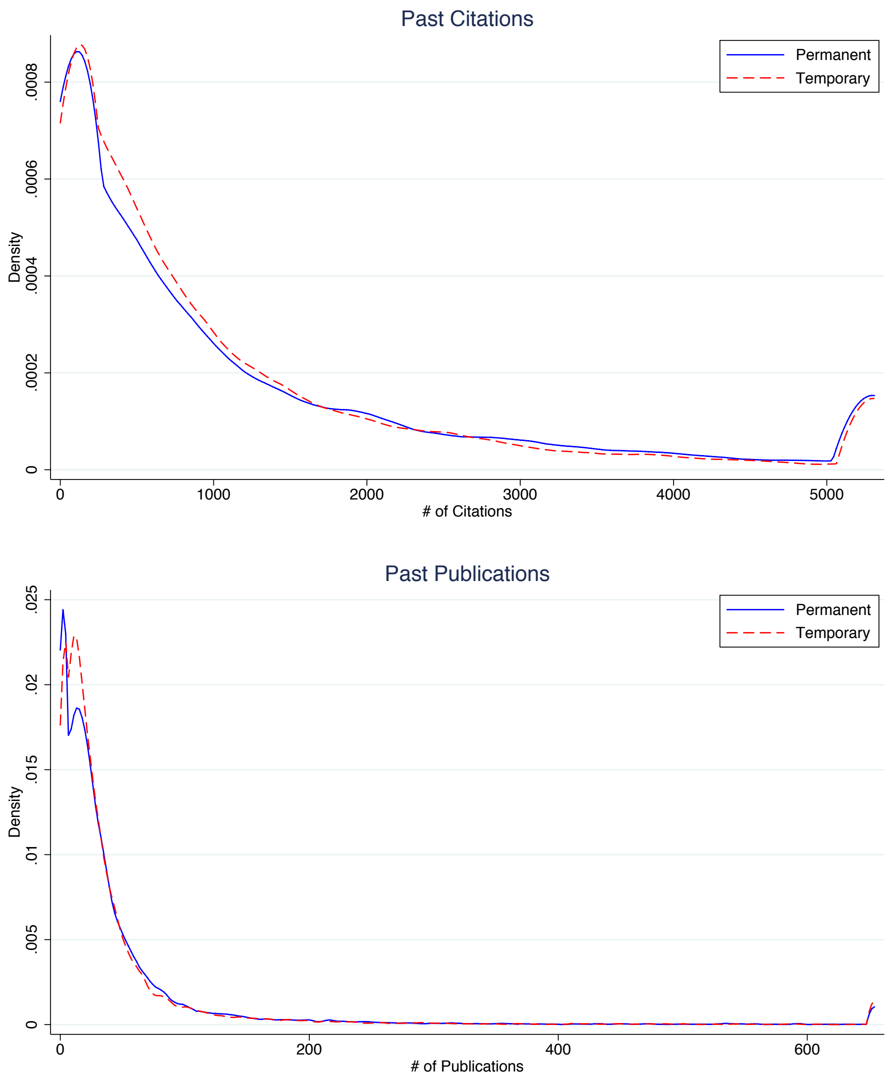
Next, Appendix Table B examines the distribution of quality by relatedness to permanent members. This is the table analogue of Appendix Figure C. The setup of this table is similar to that of Table A: I compare the difference in means and various percentiles for applicants cited by the same total number of reviewers, but by differing numbers of permanent reviewers. This table considers the distribution of quality measures for applicants cited by 0, 1, and 2 reviewers total. This encompasses 73 percent of my sample.

Overall, these results show that, among applicants cited by the same total number of reviewers, there are very few significant differences in measured quality between applicants cited by more or fewer permanent reviewers. Among applicants cited by one reviewer, there are no significant differences in citation-based quality at the mean, 1st, 5th, 10th, 25th, 50th, or 75th percentiles. At the 90th, applicants cited by one permanent reviewer have 4 more citations than applicants cited by one temporary reviewer, with this difference being significant at the 5 percent level. At the 95th percentile, this difference grows to 7 citations (but shrinks in percentage terms). Finally, at the 99th percentile, applicants cited by permanent reviewers have 10 more citations, but this difference is no longer significant.

Appendix Table B also compares the distribution of citation outcomes among applicants cited by two reviewers in total. The clear pattern that emerges here is that applicants look broadly similar but that applicants cited by two temporary reviewers appear slightly weaker than applicants cited by either one of each or two permanent members. However, it also appears that applicants cited by one of each type of reviewer have stronger records than applicants cited by two permanent members. As a result, there is no systematic relationship between number of citing permanent reviewers and application quality.

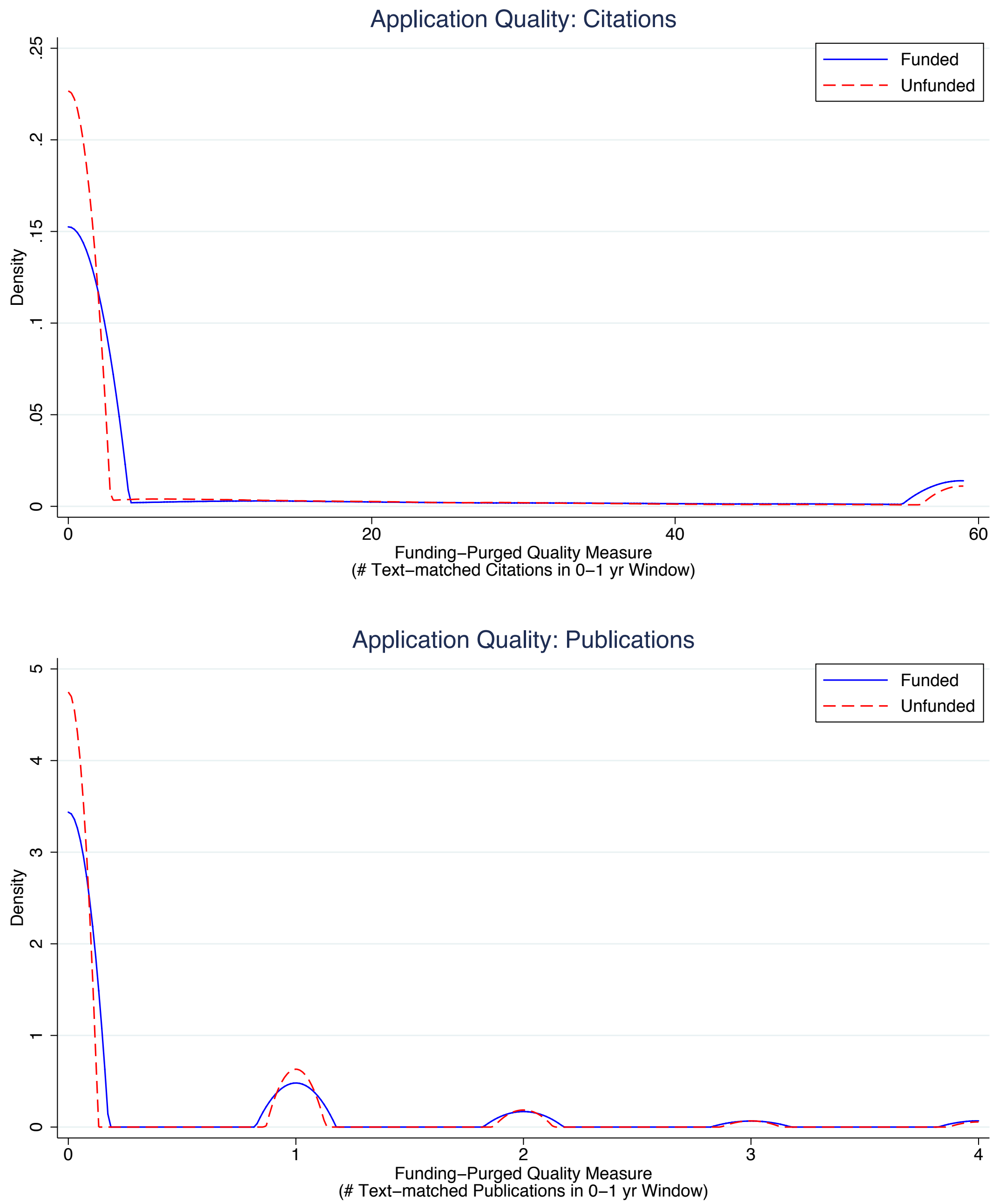
Appendix Table C provides the formal tests designed to accompany Figure 3, which shows that my measure of application quality is not impacted by a grant application's funding status. Column 1 reports the relationship between a grant's funding status and its measured quality, controlling for a linear score trend (where the impact of score is allowed to differ above and below the funding payline). Column 2 repeats this regression with finer controls for score, in this case, a quintic polynomial in score that is allowed to differ above and below the payline. In both cases, I find no significant relationship between funding and measured quality. Next, Columns 3 and 4 present IV evidence on the impact of funding on measured grant quality, where I instrument for funding with an indicator for a grant's score falling above the payline. Again I find no significant effect. Finally, Columns 5 and 6 present reduced form evidence on the relationship between falling above the payline and measured quality. I find no evidence of any association. Finally, it is worth noting that these regressions include meeting fixed effects. Within a meeting there is still variation in whether grants with the same score fall above the payline because different grants are subject to different paylines depending on what NIH Institute they are funded by. For instance, if the National Cancer Institute receives more competitive applications, than a cancer-related grant with a score of 70 may not be funded even though a diabetes-related grant with same score would be funded.

Appendix Figure A: Distribution of Past Citations and Publications, Permanent vs. Temporary Reviewers



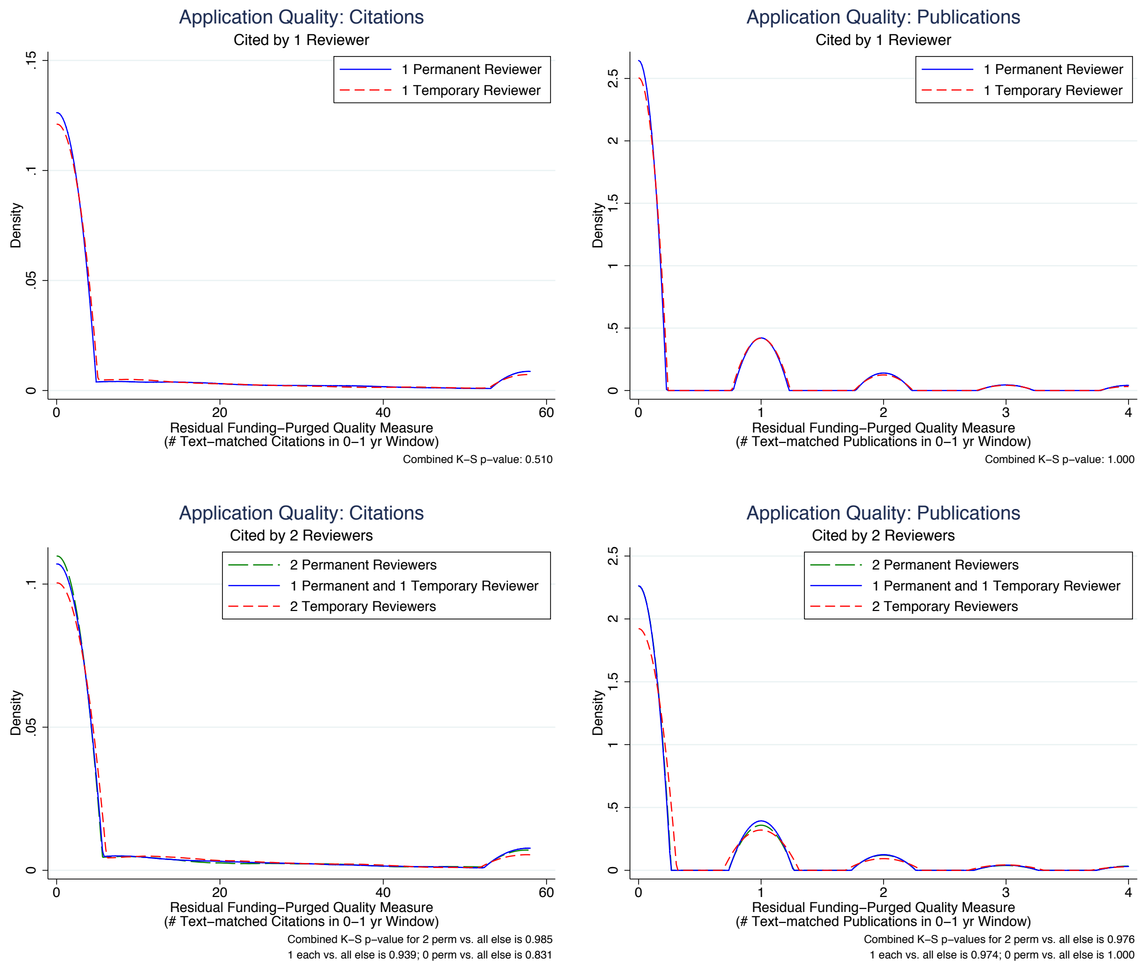
Note: Epanechnikov kernel. Publications count # of publications for which the reviewer was a first, second, or last author, published within 5 years of the relevant study section meeting. Citations count all citations to that set of publications, to 2008. Citations and Publications are top-coded at the 95th and 99th percentiles, respectively. This is done for legibility only; analyses use the full distribution of both variables. See Section II.B for additional details about how quality is constructed. A Kolmogorov-Smirnov test rejects that these two distributions are identical, with a p-value of 0.000.

Appendix Figure B: Distribution of application quality: funded and unfunded grants



Note: Epanechnikov kernel. Publications count # of text-matched publications within one year of grant review. Citations count all citations to that set of publications, to 2008. Citations and Publications are top-coded at the 95th and 99th percentiles, respectively. This is done for legibility only; analyses use the full distribution of both variables. See Section II.B and Appendix B for additional details about how quality is constructed. Kolmogorov-Smirnov tests reject that the distribution for unfunded grants is greater than for funded grants, for both publication and citation outcomes. The p-value for both tests is 0.000. It does not reject that funded grants do better on both dimensions. The p-value for both those tests is 1.000

Appendix Figure C: Application quality conditional on total # of proximate reviewers



Note: Epanechnikov kernel. Publications count # of text-matched publications within one year of grant review. Citations count all citations to that set of publications, to 2008. Citations and Publications are top-coded at the 95th and 99th percentiles, respectively. This is done for legibility only; analyses use the full distribution of both variables. See Section II.B and Appendix B for additional details about how quality is constructed. Kolmogorov-Smirnov tests reject any differences between the distributions in each figure. For the bottom panels, K-S tests are performed pairwise: distribution of those cited by 2 permanent reviewers versus distribution for those cited by less than 2 permanent reviewers; distribution for those cited by one reviewer each vs. not, distribution for those cited by 2 temporary reviewers vs. not.

Appendix Table A: Distribution of Application Quality by Funding

	Mean	*Percentiles*								
		1	5	10	25	50	75	90	95	99
Application Quality: Citations										
Funded	10.28	0	0	0	0	0	0	37	90	308
Unfunded	8.72	0	0	0	0	0	0	18	48	163
P-value	0.00	1.00	1.00	1.00	1.00	1.00	1.00	0.00	0.00	0.00
Application Quality: Publications										
Funded	3.28	0	0	0	0	0	0	1	2	3
Unfunded	2.63	0	0	0	0	0	0	1	2	4
P-value	0.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00

Notes: See notes to Table 1 for details about the sample. The quality of grant applications is measured as follows: # Publications refers to the number of research articles that the grant winner publishes in the year following the grant which share at least one salient word overlap between the grant project title and the publication title. # Citations refers to the total number of citations that accrue to this restricted set of publications, from the time of publication, to the end of my citation data in 2008. P-values for significance of percentile values are determined from a quantile regression of the quality outcome variable on an indicator variable for an application's funding status.

Appendix Table B: Distribution of Application Quality by Applicant Relatedness

	Mean	*Percentiles*								
		1	5	10	25	50	75	90	95	99
Application Quality: Citations										
No Related Reviewer										
0 Temp, 0 Perm	5.97	0	0	0	0	0	0	9	32	127
1 Related Reviewer										
1 Temp (N=7,049)	10.27	0	0	0	0	0	0	24	54	190
1 Perm (N=10,980)	11.42	0	0	0	0	0	0	28	61	200
P-value	0.16	1.00	1.00	1.00	1.00	1.00	1.00	0.03	0.07	0.61
2 Related Reviewers										
2 Temp (N=2,403)	8.72	0	0	0	0	0	0	24	50	151
1 Temp, 1 Perm (N=5,094)	11.39	0	0	0	0	0	0	28	62	213
2 Perm (N=4,841)	11.13	0	0	0	0	0	0	27	57	204
P-value*	0.08	1.00	1.00	1.00	1.00	1.00	1.00	0.67	0.24	0.04
Application Quality: Publications										
No Related Reviewer										
0 Temp, 0 Perm	0.20	0	0	0	0	0	0	1	1	3
1 Related Reviewer										
1 Temp (N=7,049)	0.32	0	0	0	0	0	0	1	2	4
1 Perm (N=10,980)	0.31	0	0	0	0	0	0	1	2	4
P-value	0.29	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
2 Related Reviewers										
2 Temp (N=2,403)	8.72	0	0	0	0	0	0	1	2	4
1 Temp, 1 Perm (N=5,094)	11.39	0	0	0	0	0	0	1	2	4
2 Perm (N=4,841)	11.13	0	0	0	0	0	0	1	2	4
P-value*	0.90	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00

Notes: See notes to Table 1 for details about the sample. The quality of grant applications is measured as follows: # Publications refers to the number of research articles that the grant winner publishes in the year following the grant which share at least one salient word overlap between the grant project title and the publication title. # Citations refers to the total number of citations that accrue to this restricted set of publications, from the time of publication, to the end of my citation data in 2008. P-values for differences in application quality for the subsample applicants cited by one reviewer only are computed from a quantile regression of the quality outcome variable on an indicator variable for an application′s funding status. P-values for differences in application quality for the subsample of applicants cited by two reviewers is computed as follows: from a quantile regression of the number of proximate permanent members. The p-value reported is the p-value on the coefficient on this linear variable, which can take values 0, 1, and 2.

Appendix Table C: Is Measured Quality Contaminated by Funding? Regression Discontinuity in Score

	Grant Application Quality (# of citations to text-matched publications within 1 year of grant review)					
	OLS		*IV*		*Reduced Form*	
	(1)	(2)	(3)	(4)	(5)	(6)
Awarded	-0.0024	0.0003	-0.0255	-1.5052		
	(0.007)	*(0.008)*	*(0.020)*	*(3.407)*		
1(Score Above Payline)					-0.0094	-0.0088
					(0.008)	*(0.018)*
Observations	99,547	99,547	99,547	99,547	99,547	99,547
R-squared	0.0673	0.0674			0.0673	0.0674
Meeting FEs	X	X	X	X	X	X
Linear Score Trends	X		X		X	
Quintics in Score		X		X		X

Notes: Coefficients are reported from a regression of grant quality on an indicator for whether the grant was funded or whether it was scored above the payline. Columns 1 and 2 examines how measured quality changes once a grant is awarded, controlling for scores in various ways. Column 1 includes the following additional controls: score and score interacted with an indicator for being above the payline. Column 2 includes quintics in score, with each score variable also interacted with an indicator for being above the payline. Columns 3 and 4 instrument awarded with being above the payline, using the same set of controls as Columns 1 and 2. Finally, Columns 5 and 6 report the reduced form regressions of measured quality on an indicator for being above the payline, using the same set of score controls as Columns 1 and 2.

*** $p<0.01$, ** $p<0.05$, * $p<0.1$

Measuring Grant Application Quality

This section describes my measure of application quality in more detail and provides additional robustness checks.

B1. Match Process

For each grant application, I have information on the name of the applicant, the title of the grant project and, in some cases, location identifiers for the applicant. I also have data from Thomson Reuters ISI Web of Science (ISI), containing information on publication titles, abstracts, and author names. To match these, I restrict to life science journal articles (e.g. excluding reviews, comments, etc.) in ISI with the same author name, published within 1 year of the study section meeting date. I have full name information in the NIH grant data, but publications are listed by last name and first and middle initial only. This results in some cases in which several authors can have the same initials (e.g. Smith, TA). In my baseline specifications, I exclude PIs with common names, defined as those last name, first initial, middle initial combinations shared by more than two individuals in PubMed. This amounts to about 7 percent of the sample being removed.

After removing common names and proceeding with an initial name and publication year match, I am left with a set of 16,134,500 possible grant-publication matches for 158,099 project titles and 3,274,225 possible publications. From this set, I compare the content of the grant project title with that of the publication title and publication abstract. I first remove a list of common stop words using the standard MySQL full test stop words list (available at http://dev.mysql.com/doc/refman/5.5/
en/fulltext-stopwords.html). After doing so, the average grant project title has 4.87 semantic words (SD 1.10). The average publication title has 8.90 words (SD 3.38); the average abstract has 52.1 words (SD 36.9). 11.58 percent of potential pairs have at least one overlapping word between the grant and publication titles. 18.08 percent of potential pairs share a common semantic word. These comparisons are made from raw words only so that "mice" and "mouse" or "males" and "male" would not match.

In our main specifications, we say that a publication and grant application are text-matched to each other if they share at least 4 semantic words in either the publication title or abstract. Consider the following example from my data.

In 1999, the National Institute of Allergy and Infectious Disease funded grant number 1R01AI045057-01 from the applicant John C Boothroyd at Stanford University. The grant project title was titled "Genetics of Invasion and Egress in Toxoplasma." This grant shows up in my raw data as follows:

Next, I search for life science publications by authors with the initials JC Boothroyd published in the first year after grant review (1999 and 2000). This yields 10 publications, of which I am excerpting five below for illustrative purposes:

Grant ID	Grant Year	Grant Title	PI Name
1R01AI045057-01	1999	Genetics of Invasion and Egress in Toxoplasma	Boothroyd, JC

Pub. ID	Pub. Year	Pub. Title	Pub. Abstract
000168366100029	2000	Ionophore-resistant mutants of Toxoplasma gondii reveal host cell permeabilization as an early event in egress	Toxoplasma gondii is an obligate intracellular pathogen within the phylum Apicomplexa. Invasion and egress by this protozoan parasite....
000165702100018	2000	Trans-spliced L30 ribosomal protein mRNA of Trypanosoma brucei is not subject to autogenous feedback control at the messenger RNA level	The regulation of gene expression in trypansomes is poorly understood but it is clear that much of this regulation, particularly of developmentally controlled genes, is post-transcriptional....
000089249600007	2000	Lytic cycle of Toxoplasma gondii	Toxoplasma gondii is an obligate intracellular pathogen within the phylum Apicomplexa. This protozoan parasite is one of the most widespread, with a broad host range including many birds and mammals and a geographic range that is nearly worldwide....
0000167020000075	2000	Toxoplasma gondii homologue of plasmodium apical membrane antigen 1 is involved in invasion of host cells	Proteins with constitutive or transient localization on the surface of Apicomplexa parasites are of particular interest for their potential role in the invasion of host cells....
000079956900015	2000	A Toxoplasma lectin-like activity specific for sulfated polysaccharides is involved in host cell infection	Toxoplasma gondii is one of the most widespread parasites of humans and animals. The parasite has a remarkable ability to invade a broad range of cells....

The first publication clearly seems related to the subject of the grant. It has 2 overlapping words in the title and 4 overlapping words in the abstract (the 4th word, "invasion," shows up later and is not reproduced here). My text matching algorithm will link this publication as related. The second publication does not seem like it has much overlap with the subject of the grant. My algorithm will not link this publication. The following three publications are more ambiguous. All of them are about "toxoplasma," which is a key word in the grant project title. The third publication only has one overlapping word ("toxoplasma") while the second has two overlapping words ("toxoplasma" and "invasion"), and the final has one overlapping word("toxoplasma") and a close second ("invasion" vs. "invade").

If we examine the list of publications actually acknowledged by the grant (this is available for funded applications only), this list includes 3 publications: the first, the third, and the fourth; the fifth publication, which looks similar in terms of word overlap, is not acknowledged. In the interest of being conservative, my main approach will match only the first publication.

B2. Robustness to alternative processes

Given the ambiguity involved in the matching process, I explore the following forms of robustness to my primary text-matching process:

1) Appendix Table D: Varying criteria for uniqueness of names
2) Appendix Table E: Varying the threshold for word overlap used to associate publications with grants
3) Appendix Tables F and G: Varying the time window for publications to be associated with grants
4) Appendix Table H: Varying the prominence of the author's contribution to a publication.
5) Appendix Table I: Compares results with alternative quality measures

Appendix Table D explores the robustness of my results to different restrictions on the types of applicant names that I include in my analysis. In my main specifications, I exclude all names with more than two individuals in PubMed who share the same last name, first and middle initial combination. The results in Appendix Table D show that my results do not change when I include all these names or when I am more restrictive, allowing only for unique last name and first and middle initial combinations.

Appendix Table E considers 8 different ways of changing threshold for how I choose whether a grant is matched to a publication. In my main specifications, I require that at least 4 semantic words be matched in either the publication title or abstract. As was discussed earlier, this may lead to cases in which publications on the same topic are missed (e.g., the third and fourth publications in the example table above.) Appendix Table B considers whether my results change when I apply different standards, both more and less stringent. Columns 1 through 4 detail results where text matching requires that X number of words overlap between the grant project title and the publication title *or* between the grant project title and the abstract, where $X = 1, 2, 3$, or 4. Because there are on average only 4.87 semantic words (SD 1.10) in the grant project title, I examine up 4 words maximum. Columns 5 through 8 repeat this exercise, but with a match defined as whether a grant project title shares X words with the publication title *and* the publication abstract (the main result is replicated in Column 5). The results show that, regardless of the exact threshold I use, my resulting estimates are similar: the impact of proximity increases with measured quality.

Appendix Tables F and G vary the time windows used to match grants to publications. Appendix Table F addresses concerns that funding may directly influence the number of citations produced by a grant by, for example, freeing up an investigator from future grant writing so that he can concentrate on research. Instead of including articles published after the grant is reviewed, Appendix Table F restricts my analysis to articles published one year before a grant is reviewed. These publications are highly likely to be based off research that existed before the grant was reviewed, but cannot have been influenced by the grant funds. Using this metric, I find nearly identical measures of bias and information. Appendix Table G addresses the opposite concern, that a one-year window after review may be insufficient to assess the quality of grant applications. Instead, I use a five year window following review and find that my results are both qualitatively and quantitatively similar. My estimates are very similar.

Finally, the next set of results explores the validity of my quality measures more broadly. The goal of my quality measures is to capture the quality of the research written into the grant application at the time of grant review. One possible concern with examining all publications by an author is that some of these publications may be ones for which the author made few substantive intellectual contributions, and which might not reflect his or her research program. Life science articles often have many authors and collaborators on a project may receive authorial credit for minor contributions such as sharing equipment or making figures. To address this, Appendix Table H restricts my match process to publications for which the grant applicant was the first, second, or last author. In the life sciences, contributions can be inferred from authorship position with earlier authors deserving more credit, and the last author being the primary investigator. Again, I find that the impact of proximity increases in application quality.

Finally, Appendix Table I shows that my results are robust to splitting my sample based on various non-residualized measures of quality: whether or not an application goes on to produce any citations to text-matched publications within the first year at all; those that produce publications cited at the 95th percentile of its field-year cohort vs. not; and those that produce publications cited at the 99th percentile of this distribution vs not. In all these cases, I find a stronger effect of proximity on higher quality applications.[21] For example, among applications that go on to produce publications in the top 99th percentile of their cohort's citation distribution, each additional reviewer increases their likelihood of funding by 2.1 percentage points, from a baseline funding rate for that group of 28.2 percent, or a 7.4 percent increase. This is similar to the magnitude I find for my top quartile applications from Table 7.

[21] It is not possible to explore the impact of proximity on the funding outcomes of particularly low quality candidates according to unresidualized measures of quality. This is because there is significant bunching of applications at zero publications and citations.

Appendix Table D: What is the impact of proximity by application quality? Robustness to alternative name-frequencies

	Quartiles of Residual Application Quality				
	All	*Bottom*	*Second*	*Third*	*Top*
	(1)	(2)	(3)	(4)	(5)
	Dependent Variable: 1(Score Above Payline)				
# Proximate Permanent Reviewers	0.0047**	-0.0018	-0.0007	0.0034	0.0177***
	(0.002)	*(0.004)*	*(0.004)*	*(0.005)*	*(0.005)*
Observations	86,486	20,775	22,004	21,663	22,044
R-squared	0.0694	0.1535	0.1293	0.1186	0.1330
	Dependent Variable: Score				
# Proximate Permanent Reviewers	0.1763*	-0.1014	0.0478	0.2535	0.7705***
	(0.099)	*(0.181)*	*(0.231)*	*(0.329)*	*(0.230)*
Observations	53,183	14,942	13,327	11,176	13,738
R-squared	0.1248	0.2081	0.2145	0.2597	0.2068
	Dependent Variable: 1(Scored at all)				
# Proximate Permanent Reviewers	0.0014	-0.0005	-0.0097**	0.0026	0.0104**
	(0.002)	*(0.004)*	*(0.005)*	*(0.007)*	*(0.005)*
Observations	86,486	20,775	22,004	21,663	22,044
R-squared	0.0911	0.1549	0.1463	0.1352	0.1399
Meeting FEs	X	X	X	X	X
# of Proximate Reviewer FEs	X	X	X	X	X

Notes: This table presents the same results as Table 7 but restricting to investigators who have a unique last name, first and middle initial combination in PubMed. Coefficients are reported from a regression of committee decisions (above payline, score, or scored at all) on # of proximate permanent reviewers, controlling for meeting level fixed effects and fixed effects for total proximity. Panel 1 regressions use the same specification as Column 2 in Table 5; Panel 2 uses the same specification as Column 5 in Table 5; the final panel uses the same specification as Column 8. Columns 2 through 5 split the sample based on quartiles of residual application quality. To calculate this, I regress application quality in citations on dummies for female, Hispanic, east Asian, south Asian, M.D., Ph.D., fixed effects for decile bins for both past publication and citations, and fixed effects for number of past R01 and other grants, and taking the residuals from this regression.

*** p<0.01, ** p<0.05, * p<0.1

Appendix Table E: What is the impact of proximity by application quality? Robustness to alternative text-matching word thresholds

	Dependent Variable: 1(Score Above Payline)							
	Quartiles of Residual Application Quality							
	Bottom	*Second*	*Third*	*Top*	*Bottom*	*Second*	*Third*	*Top*
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	>4 Overlapping Words Title OR Abstract				*>4 Overlapping Words Title AND Abstract*			
# Proximate Permanent Reviewers	-0.0006	0.0000	0.0039	0.0144***	0.0004	0.0015	0.0042	0.0083
	(0.004)	*(0.004)*	*(0.005)*	*(0.005)*	*(0.004)*	*(0.004)*	*(0.005)*	*(0.006)*
	>3 Overlapping Words Title OR Abstract				*>3 Overlapping Words Title AND Abstract*			
# Proximate Permanent Reviewers	-0.0001	0.0012	0.0085	0.0085*	-0.0009	0.0036	0.0048	0.0120**
	(0.004)	*(0.004)*	*(0.006)*	*(0.004)*	*(0.004)*	*(0.004)*	*(0.005)*	*(0.005)*
	>2 Overlapping Words Title OR Abstract				*>2 Overlapping Words Title AND Abstract*			
# Proximate Permanent Reviewers	-0.0011	-0.0007	0.0120**	0.0085*	-0.0019	0.0050	0.0082	0.0077*
	(0.004)	*(0.004)*	*(0.006)*	*(0.004)*	*(0.004)*	*(0.004)*	*(0.005)*	*(0.004)*
	>1 Overlapping Words Title OR Abstract				*>1 Overlapping Words Title AND Abstract*			
# Proximate Permanent Reviewers	-0.0013	0.0079*	0.0077	0.0033	0.0009	0.0010	0.0032	0.0100**
	(0.004)	*(0.004)*	*(0.006)*	*(0.005)*	*(0.004)*	*(0.004)*	*(0.006)*	*(0.004)*
Observations	23,854	24,368	23,458	21,878	23,218	24,101	23,602	22,637
Meeting FEs	X	X	X	X	X	X	X	X
# of Proximate Reviewer FEs	X	X	X	X	X	X	X	X

Notes: These regressions repeat Columns 2-5 from the first panel of Table 7. Coefficients are reported from a regression of 1(score above payline) on # of proximate permanent reviewers, controlling for meeting level fixed effects and fixed effects for total proximity. Columns 1 through 4 split the sample based on quartiles of residual application quality, where applicant quality is defined based on text matching that requires X words of overlap between the grant project title and the title of the publication or its abstract, where X = 1, 2, 3, 4. Columns 5-8 repeat this same exercise, except using text matching that requires word overlap with both the publication title and its abstract.

*** p<0.01, ** p<0.05, * p<0.1

Appendix Table F: What is the impact of proximity by application quality? Grant quality measured from articles published 1 year before grant review

	Quartiles of Residual Application Quality (Citations to Publications 1 Year Before Meeting)				
	All	*Bottom*	*Second*	*Third*	*Top*
	(1)	(2)	(3)	(4)	(5)
	Dependent Variable: 1(Score Above Payline)				
# Proximate Permanent Reviewers	0.0050**	-0.0018	0.0024	0.0044	0.0193***
	(0.002)	*(0.004)*	*(0.004)*	*(0.004)*	*(0.005)*
Observations	93,558	22,367	23,925	23,753	23,513
R-squared	0.0688	0.1492	0.1266	0.1082	0.1279
	Dependent Variable: Score				
# Proximate Permanent Reviewers	0.1641*	-0.0992	0.0808	0.2468	0.8997***
	(0.094)	*(0.165)*	*(0.202)*	*(0.298)*	*(0.258)*
Observations	57,613	16,336	15,037	12,327	13,913
R-squared	0.1224	0.1957	0.2043	0.2432	0.2104
	Dependent Variable: 1(Scored at all)				
# Proximate Permanent Reviewers	0.0012	-0.0013	-0.0054	-0.0046	0.0073
	(0.002)	*(0.003)*	*(0.004)*	*(0.006)*	*(0.006)*
Observations	93,558	22,367	23,925	23,753	23,513
R-squared	0.0899	0.1482	0.1395	0.1313	0.1383
Meeting FEs	X	X	X	X	X
# of Proximate Reviewer FEs	X	X	X	X	X

Notes: See notes to Table 1 for details about the sample. The key difference between this table and Table 7 is that application quality is measured using publications text-matched to the grant application project title in the one year before the grant meeting. Coefficients are reported from a regression of committee decisions (above payline, score, or scored at all) on # of proximate permanent reviewers, controlling for meeting level fixed effects and fixed effects for total proximity. Panel 1 regressions use the same specification as Column 2 in Table 5; Panel 2 uses the same specification as Column 5 in Table 5; the final panel uses the same specification as Column 8. Columns 2 through 5 split the sample based on quartiles of residual application quality. To calculate this, I regress application quality in citations on dummies for female, Hispanic, east Asian, south Asian, M.D., Ph.D., fixed effects for decile bins for both past publication and citations, and fixed effects for number of past R01 and other grants, and taking the residuals from this regression.

*** $p<0.01$, ** $p<0.05$, * $p<0.1$

Appendix Table G: What is the impact of proximity by application quality? Grant quality measured from articles published 0-5 years after grant review

	Quartiles of Residual Application Quality (Citations to Publications up to 5 Years After Meeting)				
	All	*Bottom*	*Second*	*Third*	*Top*
	(1)	(2)	(3)	(4)	(5)
	Dependent Variable: 1(Score Above Payline)				
# Proximate Permanent Reviewers	0.0050**	-0.0021	-0.0010	0.0069	0.0172***
	(0.002)	*(0.004)*	*(0.004)*	*(0.005)*	*(0.005)*
Observations	93,558	22,764	23,981	23,794	23,019
R-squared	0.0688	0.1495	0.1206	0.1101	0.1401
	Dependent Variable: Score				
# Proximate Permanent Reviewers	0.1641*	-0.1145	-0.1030	0.3783	0.5820**
	(0.094)	*(0.168)*	*(0.216)*	*(0.306)*	*(0.234)*
Observations	57,613	15,954	14,342	12,440	14,877
R-squared	0.1224	0.2032	0.2067	0.2382	0.2052
	Dependent Variable: 1(Scored at all)				
# Proximate Permanent Reviewers	0.0012	-0.0013	-0.0066	0.0052	0.0095*
	(0.002)	*(0.003)*	*(0.005)*	*(0.006)*	*(0.005)*
Observations	93,558	22,764	23,981	23,794	23,019
R-squared	0.0899	0.1504	0.1403	0.1275	0.1426
Meeting FEs	X	X	X	X	X
# of Proximate Reviewer FEs	X	X	X	X	X

Notes: See notes to Table 1 for details about the sample. The key difference between this table and Table 7 is that application quality is measured using publications text-matched to the grant application project title up to five years after the grant meeting. Coefficients are reported from a regression of committee decisions (above payline, score, or scored at all) on # of proximate permanent reviewers, controlling for meeting level fixed effects and fixed effects for total proximity. Panel 1 regressions use the same specification as Column 2 in Table 5; Panel 2 uses the same specification as Column 5 in Table 5; the final panel uses the same specification as Column 8. Columns 2 through 5 split the sample based on quartiles of residual application quality. To calculate this, I regress application quality in citations on dummies for female, Hispanic, east Asian, south Asian, M.D., Ph.D., fixed effects for decile bins for both past publication and citations, and fixed effects for number of past R01 and other grants, and taking the residuals from this regression.

*** p<0.01, ** p<0.05, * p<0.1

Appendix Table H: What is the impact of proximity by application quality? Grant quality measured from first, second, and last authorship position articles

	Quartiles of Residual Application Quality				
	All	*Bottom*	*Second*	*Third*	*Top*
	(1)	(2)	(3)	(4)	(5)
	Dependent Variable: 1(Score Above Payline)				
# Proximate Permanent Reviewers, (based on citations by first, second, and last authors only)	0.0083***	0.0060	0.0031	-0.0008	0.0160**
	(0.003)	*(0.005)*	*(0.006)*	*(0.007)*	*(0.006)*
Observations	93,558	22,463	23,929	23,360	23,806
R-squared	0.0678	0.1504	0.1249	0.1102	0.1257
	Dependent Variable: Score				
# Proximate Permanent Reviewers, (based on citations by first, second, and last authors only)	0.2423*	0.1487	0.2073	-0.2040	0.5668*
	(0.130)	*(0.227)*	*(0.311)*	*(0.417)*	*(0.311)*
Observations	57,613	16,081	14,593	12,056	14,883
R-squared	0.1218	0.1999	0.2077	0.2433	0.1967
	Dependent Variable: 1(Scored at all)				
# Proximate Permanent Reviewers, (based on citations by first, second, and last authors only)	0.0032	-0.0010	-0.0099*	0.0060	0.0148**
	(0.003)	*(0.004)*	*(0.006)*	*(0.010)*	*(0.006)*
Observations	93,558	22,463	23,929	23,360	23,806
R-squared	0.0866	0.1480	0.1392	0.1293	0.1299
Meeting FEs	X	X	X	X	X
# of Proximate Reviewer FEs	X	X	X	X	X

Notes: See notes to Table 1 for details about the sample. The key difference between this table and Table 7 is that proximity is determined by citations made to the applicant's work, only by reviewers who were the first, second, or last authors on the citing publication. Coefficients are reported from a regression of committee decisions (above payline, score, or scored at all) on # of proximate permanent reviewers, controlling for meeting level fixed effects and fixed effects for total proximity. Panel 1 regressions use the same specification as Column 2 in Table 5; Panel 2 uses the same specification as Column 5 in Table 5; the final panel uses the same specification as Column 8. Columns 2 through 5 split the sample based on quartiles of residual application quality. To calculate this, I regress application quality in citations on dummies for female, Hispanic, east Asian, south Asian, M.D., Ph.D., fixed effects for decile bins for both past publication and citations, and fixed effects for number of past R01 and other grants, and taking the residuals from this regression.

*** p<0.01, ** p<0.05, * p<0.1

Appendix Table I: What is the impact of proximity by application quality? Alternative Quality Meaures

	Alternative Measures of Application Quality					
	Any Citations?		Top 99th Percentile Pubs		Top 99th Percentile Pubs	
	No	*Yes*	*No*	*Yes*	*No*	*Yes*
	(1)	(2)	(3)	(4)	(3)	(4)
	Dependent Variable: 1(Score Above Payline)					
# Proximate Permanent Reviewers	0.0036*	0.0133**	0.0035*	0.0207***	0.0035*	0.0207***
	(0.002)	*(0.006)*	*(0.002)*	*(0.008)*	*(0.002)*	*(0.008)*
Observations	80,379	13,179	85,946	7,612	85,946	7,612
R-squared	0.0961	0.2292	0.0940	0.3148	0.0940	0.3148
	Dependent Variable: Score					
# Proximate Permanent Reviewers	0.1249	0.5062**	0.1327	0.4285	0.1327	0.4285
	(0.105)	*(0.249)*	*(0.101)*	*(0.353)*	*(0.101)*	*(0.353)*
Observations	48,435	9,178	51,995	5,618	51,995	5,618
R-squared	0.1522	0.2798	0.1478	0.3677	0.1478	0.3677
	Dependent Variable: 1(Scored at All)					
# Proximate Permanent Reviewers	0.0006	0.0041	0.0008	0.0084	0.0008	0.0084
	(0.002)	*(0.006)*	*(0.002)*	*(0.007)*	*(0.002)*	*(0.007)*
Observations	80,379	13,179	85,946	7,612	85,946	7,612
R-squared	0.1360	0.2326	0.1342	0.3050	0.1342	0.3050
Meeting FEs	X	X	X	X	X	X
# of Proximate Reviewer FEs	X	X	X	X	X	X
Past Performance, Past Grants, and Demographics	X	X	X	X	X	X

Notes: See notes to Table 1 for details about the sample. Coefficients are reported from a regression of committee decisions (above payline, score, or scored at all) on # of proximate permanent reviewers, controlling for meeting level fixed effects and fixed effects for total proximity. Each column presents the main regression from Table 5 on a diferent sample based on publication outcomes related to the application. Columns 1 and 2 compare applications with any citations to text-matched publications within on year of grant review, versus those without. Columns 3 and 4 compare applications that then go on to produce a text-matched publication within one year of grant review, where that publication is cited at the top 95th percentile of all publications from the same year, based on citations in 2008 -- versus not. Columns 5 and 6 do the same for publications at the 99th percentile of citations. For all these regressions, I include controls for applicant characteristics: female, Hispanic, east Asian, south Asian, M.D., Ph.D., fixed effects for decile bins for both past publication and citations, and indicators for the number of past R01 and other NIH grants an applicant has won, as well as indicators for how many she has applied to.

*** p<0.01, ** p<0.05, * p<0.1

Established vs. New Investigators

Appendix Table J examines how my main estimates in Table 7 vary by whether or not a grant is new (Columns 1-5), and whether or not an investigator is new (Columns 6-10). My results indicate that the impact of proximity is similar for new and renewal grants when viewed in percentage point terms. However, new grants have a lower average probability of being funded, relative to renewal grants. Among new applications in the top quartile of quality, an additional related reviewer increases the application's likelihood of funding by 1.9 percentage points from a base of 17.7 percent, or a 10.1 percent increase. Among new applications in the bottom quartile of quality, proximity decreases an application's likelihood of funding by 0.8 percentage points, from a base of 21.1 percent, or a 4.8 percent decrease. This suggests that the informational advantage of related reviewers may be greater for new applications.

My results are noisier for entirely new investigators. While I cannot reject that the effect for new investigators is similar to my estimates for established investigators, I cannot reject that they are zero either. If it is the case that I find stronger effects of relatedness for established investigators, this may be because reviewers are familiar with the research agendas of established investigators. While most new investigators have a history of publications, these articles are often written in conjunction with a more senior scientist who funds the research. When reviewers cite publications by new investigators, they may be more familiar with the work of the senior scientist, rather than that of the new investigator herself, who is likely to have been a graduate student, postdoc, or unfunded junior academic at the time. As a result, proximity to new investigators may convey less information that proximity to established scientists.

Appendix Table J: Impact of Proximity, Heterogeneity by Grant and Applicant Type

	Dependent Variable: 1(Score Above Payline)									
	Quartiles of Residual Application Quality					**Quartiles of Residual Application Quality**				
	All	*Bottom*	*Second*	*Third*	*Top*	*All*	*Bottom*	*Second*	*Third*	*Top*
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Renewal Grants					*Established Investigators*				
# Proximate Permanent Reviewers	0.0066*	-0.0032	-0.0048	0.0151	0.0187*	0.0052**	-0.0009	0.0030	0.0025	0.0154***
	(0.004)	*(0.009)*	*(0.009)*	*(0.011)*	*(0.010)*	*(0.002)*	*(0.004)*	*(0.004)*	*(0.005)*	*(0.005)*
Observations	27,782	6,530	6,992	7,155	7,105	84,463	20,166	21,444	21,445	21,408
R-squared	0.1385	0.3499	0.3152	0.2791	0.2977	0.0713	0.1608	0.1340	0.1169	0.1360
	New Grants					*New Investigators*				
# Proximate Permanent Reviewers	0.0030	-0.0083*	0.0066	-0.0072	0.0189***	-0.0031	0.0044	-0.0223	0.0054	0.0021
	(0.002)	*(0.005)*	*(0.005)*	*(0.006)*	*(0.006)*	*(0.009)*	*(0.024)*	*(0.031)*	*(0.041)*	*(0.033)*
Observations	65,776	15,935	16,661	16,639	16,541	9,095	2,271	2,282	2,181	2,361
R-squared	0.0644	0.1606	0.1443	0.1323	0.1558	0.2465	0.5362	0.5743	0.5400	0.5989
Meeting FEs	X	X	X	X	X	X	X	X	X	X
# of Proximate Reviewer FEs	X	X	X	X	X	X	X	X	X	X

Notes: See notes to Table 1 for details about the sample, and Table 7 for details about the regression specification. In all regressions, the dependent variable is an indicator for score being greater than the funding payline. Columns 1-5 split the sample based on whether a grant application is a renewal application (top panel) or whether it was for a new project (bottom panel). Column 1 examines the full sample of the relevant set of grants. Columns 2 through 5 split the sample based on quartiles of residual application quality. To calculate this, I regress application quality in citations on dummies for female, Hispanic, east Asian, south Asian, M.D., Ph.D., fixed effects for decile bins for both past publication and citations, and fixed effects for number of past R01 and other grants, and taking the residuals from this regression. Columns 6-10 repeat this exercise for the sample of established investigators (those who have had prior NIH funding as a primarily investigator of any kind) and new investigators, those who have not.

*** p<0.01, ** p<0.05, * p<0.1

Alternative Identification Strategy

In my main specification, I identify the effect of proximity to more influential reviewers (permanent vs. temporary). This approach relies on the assumption that controlling for the total number of reviewers who cite an applicant is an adequate control for unobserved differences in quality that may be correlated with whether an applicant is cited by a permanent reviewer. A different approach would be to use applicant fixed effects to control for quality, compare the funding outcomes of applications from the *same* applicant across meetings in which the applicant is cited by different total numbers of reviewers.[22]

The downside of this approach is that applicant fixed effects only control for time-invariant unobserved quality. If there are aspects of the quality of an applicant's proposal that are not controlled for with information on past publications and grant histories, then this may bias my results.

This second approach also captures a slightly different causal effect: the effect of being related to an additional reviewer, as opposed to being related to a more influential reviewer. The relative magnitudes of these effects are theoretically ambiguous: if only permanent reviewers have influence, then the effect of being related to a permanent reviewer (conditional on total proximity) will be larger than the effect of being related to an additional member (because that additional member may be temporary and thus, in this example, inconsequential). If, on the other hand, temporary members have as much influence as permanent ones, then the composition of related reviewers would not matter, but the number would.

Appendix Table K reports estimates from this alternative identification strategy. My results are similar. Overall, I find a significant impact of proximity on an applicant's likelihood of funding, the score that it receives, and the likelihood that it is scored at all. This effect is largely increasing with the quality of the application, although my estimates peak at the 3rd quartile rather than the 4th. For example, I find that each additional proximate reviewer—either temporary or permanent—increases an applicant's likelihood of funding by 0.61 percentage points or 2.9 percent. For the top quartile, this effect rises to 0.87 percentage points or 3.9 percent.

[22] In my alternative specification using applicant fixed effects, the analogous regression equation is given by:

$$\text{Assessment}_{icmt} = a_0 + a_1 \text{Total Proximity}_{icmt} + \mu X_{icmt} + \delta_i + \varepsilon_{icmt}.$$

Appendix Table K: What is the impact of proximity by application quality? Applicant Fixed Effects

	Quartiles of Residual Application Quality				
	All	*Bottom*	*Second*	*Third*	*Top*
	(1)	(2)	(3)	(4)	(5)
	Dependent Variable: 1(Score Above Payline)				
# of Proximate Reviewers	0.0061***	0.0039**	0.0090***	0.0111***	0.0087**
	(0.001)	*(0.002)*	*(0.003)*	*(0.004)*	*(0.004)*
Observations	93,558	22,463	23,929	23,360	23,806
R-squared	0.4527	0.5296	0.6125	0.6379	0.6276
	Dependent Variable: Score				
# of Proximate Reviewers	0.2679***	0.1994**	0.2686	0.7719**	0.4960**
	(0.058)	*(0.087)*	*(0.189)*	*(0.364)*	*(0.196)*
Observations	57,613	16,081	14,593	12,056	14,883
R-squared	0.5452	0.5993	0.6937	0.7306	0.7055
	Dependent Variable: 1(Score above payline)				
# of Proximate Reviewers	0.0111***	0.0088***	0.0144***	0.0169***	0.0118***
	(0.001)	*(0.002)*	*(0.003)*	*(0.005)*	*(0.004)*
Observations	93,558	22,463	23,929	23,360	23,806
R-squared	0.5636	0.5948	0.6780	0.7120	0.7135
Applicant FEs	X	X	X	X	X
Past Performance, Past Grants	X	X	X	X	X

Notes: See notes to Table 7 for details about the sample and variable construction. Coefficients are reported from a regression of committee decisions (above payline, score, or scored at all) on # of proximate reviewers (either permanent or temporary), controlling for applicant fixed effects and other applicant characteristics. Column 1 estimates this regression on the whole sample. Columns 2 through 5 split the sample based on quartiles of residual application quality. To calculate this, I regress application quality in citations on dummies for female, Hispanic, east Asian, south Asian, M.D., Ph.D., fixed effects for decile bins for both past publication and citations, and fixed effects for number of past R01 and other grants, and taking the residuals from this regression. For all these regressions, I include applicant fixed effects as well as controls for time varying applicant characteristics: fixed effects for decile bins for both past publication and citations, and indicators for the number of past R01 and other NIH grants an applicant has won, as well as indicators for how many she has applied to.

*** $p<0.01$, ** $p<0.05$, * $p<0.1$

Additional robustness checks

This section provides broader tests of my empirical specifications.

A key identifying assumption is that my measure of quality is not affected by whether individuals are actually funded. Figure 3 provides my primary evidence that this is the case. Another test of my assumption that citations are not directly affected by funding is to ask whether I find bias in the review of inframarginal grants, that is grants that are well above or well below the funding margin. All grants in either group have the same funding status so any bias estimate cannot be attributed to differences in funding. Because I hold funding status constant, I can only assess the impact that related permanent members have on an applicant's score not on an applicant's funding status. Appendix Table L reports these results. The top panel reports the impact of proximity on scores, using funded grants only. The bottom panel does the same for unfunded grants. In both cases, I find an increasing effect of proximity with quality. The magnitudes are somewhat smaller than in my main regression; because these are subsamples, there is no reason to expect that the magnitude of the effect of proximity to be the same as it is for the entire sample.

Another potential concern is that committees may defy instructions and evaluate grant applications not on the basis of the specific research in the proposal, but on the quality of projects that reviewers suspect the grant funding may cross subsidize. In this case, by using text-matching to restrict my main quality measure to be based on articles that are closely related to the grant proposal topic, I am potentially missing other research that reviewers might be anticipating when they evaluate a grant proposal. To test whether this is the case, I use grant acknowledgement data recorded in the National Library of Medicine's PubMed database to match funded grants to all the articles that it produces, regardless of topic or date of publication. Because this process requires that a grant application actually be funded, I am only able to examine the impact of proximity on scores, rather than on funding likelihood or the likelihood of being scored. For the set of funded grants, Appendix Table M reruns my core regressions using citations to publications that explicitly acknowledge a grant as my measure of quality, and scores as my outcome measure. I find results that are consistent with my primary findings, though of a slightly smaller magnitude.[23]

Finally, despite the tests presented in Tables 4 and Appendix Figure C, there may still potentially be a correlation between relatedness to permanent members and unobserved aspects of applicant quality. If this were driving my results, one might expect the impact of relatedness I find in Table 7 to appear similar to the impact of observed measures of quality, insofar as observed and unobserved quality may be correlated. Appendix Table N shows that this is not the case by comparing the effect of past citations on funding probability, by application quality, to the effect of relatedness. In order to implement this comparison

[23]This analysis differs slightly from my main results using citations because general citations cannot be computed for publications in PubMed. A limited set of citations can, however, be computed using publications in PubMed Central (PMC). PMC contains a subset of life sciences publications made available for free. While this is not as comprehensive a universe as that of Web of Science, it contains, for recent years, all publications supported by NIH dollars. Undercounting of publications would, further, not bias my result as long as it does not vary systematically by whether an applicant is related to a permanent or to a temporary member.

directly, the specifications in Appendix Table N differ in a two ways from Table 7. First, recall that application quality quartiles used in Table 7 are defined in terms of residual quality, accounting for demographics, grant history, and past publications. In order to examine the impact of past citations on funding outcomes, I now residualize quality using all variables except publication history. Second, the specification used in Table 7 control for fixed effects in deciles of past publications and citations. In order to clearly assess the impact of past citations, Appendix Table N includes only controls for meeting fixed effects and fixed effects for the number of total related reviewers. The top panel of Appendix Table N shows that the marginal impact of an applicant's past citations is the same across quality quartiles. By contrast, the findings in the bottom panel show that the impact of relatedness is zero for the lowest quality applicants and increasing in quality thereafter.

Appendix Table L: What is the impact of proximity by application quality? Inframarginal grant applications

		Quartiles of Residual Application Quality			
	All	*Bottom*	*Second*	*Third*	*Top*
	(1)	(2)	(3)	(4)	(5)
			Funded Sample: Score		
# Proximate Permanent Reviewers	0.1512*	0.1445	0.0642	0.0346	0.3001
	(0.078)	*(0.174)*	*(0.193)*	*(0.197)*	*(0.226)*
Observations	24,395	5,889	6,076	6,278	6,152
R-squared	0.1613	0.3622	0.3771	0.3773	0.3098
			Unfunded Sample: Score		
# Proximate Permanent Reviewers	0.1012	-0.0222	0.3286	0.1348	0.1694
	(0.091)	*(0.183)*	*(0.261)*	*(0.335)*	*(0.175)*
Observations	33,218	8,719	7,658	6,554	10,287
R-squared	0.1786	0.3090	0.3492	0.3785	0.2495
Meeting FEs	X	X	X	X	X
# of Proximate Reviewer FEs	X	X	X	X	X

Notes: Sample is the set of funded grants and of unfunded grants, treated separately. Coefficients are reported from a regression of score on # of proximate permanent reviewers, controlling for meeting level fixed effects and fixed effects for total proximity. Column 1 examines all applications. Columns 2 through 5 split the sample based on quartiles of residual application quality, within the set of funded or unfunded grants. To calculate this, I regress application quality in citations on dummies for female, Hispanic, east Asian, south Asian, M.D., Ph.D., fixed effects for decile bins for both past publication and citations, and fixed effects for number of past R01 and other grants, and taking the residuals from this regression.

*** $p<0.01$, ** $p<0.05$, * $p<0.1$

Appendix Table M: What is the impact of proximity by application quality? Explicit grant acknowledgements for the sample of funded grants

	Quartiles of Residual Application Quality (based on explicit grant acknowledgements)				
	All	*Bottom*	*Second*	*Third*	*Top*
	(1)	(2)	(3)	(4)	(5)
			Dependent Variable: Score		
# Proximate Permanent Reviewers	0.1512*	0.1445	0.0642	0.0346	0.3001
	(0.078)	*(0.174)*	*(0.193)*	*(0.197)*	*(0.226)*
Observations	24,395	5,889	6,076	6,278	6,152
R-squared	0.1613	0.3622	0.3771	0.3773	0.3098
Meeting FEs	X	X	X	X	X
# of Proximate Reviewer FEs	X	X	X	X	X

Notes: Sample is funded grants only. Coefficients are reported from a regression of committee score on # of proximate reviewers, controlling for meeting level fixed effects and fixed effects for total proximity. Proximity to permanent reviewers is defined as the number of permanent reviewers who have cited the applicant's research in the 5 years prior to grant review. "Grant Application Quality" is defined as the number of citations up to 2008, for all publications that explicitly acknowledge funding from a grant, in the 100s unit. Columns 2 through 5 split the sample based on quartiles of residual application quality. To calculate this, I regress application quality in citations (using explicit acknowledgements) on dummies for female, Hispanic, east Asian, south Asian, M.D., Ph.D., fixed effects for decile bins for both past publication and citations, and fixed effects for number of past R01 and other grants, and taking the residuals from this regression. Dependent varialbe is the number of citations to publications that explicitly acknowledge the funded grant.

*** $p<0.01$, ** $p<0.05$, * $p<0.1$

Appendix Table N: Are estimates likely to be driven by unobserved quality?
Impact of past citations on funding probability

	Quartiles of Residual Application Quality				
	All	*Bottom*	*Second*	*Third*	*Top*
	(1)	(2)	(3)	(4)	(5)
	Dependent Variable: 1(Score Above Payline)				
# Past Citations (100s)	0.0018***	0.0017***	0.0019***	0.0021***	0.0016***
	(0.000)	*(0.000)*	*(0.000)*	*(0.000)*	*(0.000)*
Observations	93,558	23,105	23,756	23,945	22,752
R-squared	0.0719	0.1363	0.1410	0.1176	0.1468
# Proximate Permanent Reviewers	0.0050**	-0.0006	0.0036	0.0082*	0.0090**
	(0.002)	*(0.004)*	*(0.004)*	*(0.005)*	*(0.004)*
Observations	93,558	23,105	23,756	23,945	22,752
R-squared	0.0688	0.1335	0.1379	0.1150	0.1441
Meeting FEs	X	X	X	X	X
# of Proximate Reviewer FEs	X	X	X	X	X

Notes: See notes to Table 1 for details about the sample. In the first panel, coefficients are reported from a regression of committee decisions (above payline, score, or scored at all) on # of past citations, controlling for meeting level fixed effects and fixed effects for total proximity. Panel 2 does the same but with # of proximate permanent reviewers, controlling for meeting level fixed effects and fixed effects for total proximity. Columns 2 through 5 split the sample based on quartiles of residual application quality. To calculate this, I regress application quality in citations on dummies for female, Hispanic, east Asian, south Asian, M.D., Ph.D., and fixed effects for number of past R01 and other grants, and taking the residuals from this regression. For this specification, these residuals are calculated without information on past publications and citations.

*** $p<0.01$, ** $p<0.05$, * $p<0.1$

Theoretical Model and Alternative Estimation Strategy

In my final set of appendices, I present a model of expertise and bias in decision-making that can be directly estimated using a linear analysis in my data. The benefit of this approach, relative to the approach in the main body of the paper, is that it allows me to: 1) formally define expertise and bias and 2) show how these unobserved parameters and signals impact the equilibrium relationship between observable funding decisions, proximity to applicants, and realized grant quality; and 3) show how these parameters can be recovered from a linear regression of committee decisions on relatedness and quality.

F1. Model

A grant application has some true quality Q^* and, if approved, the committee receives a payoff of Q^*. If the application is rejected, the committee receives its outside option U, where $U > E(Q^*)$. Applications either work in the same area as the reviewer ("proximate," given by $P = 1$) or not ($P = 0$). This model makes the simplifying assumption that committees can observe whether an application is related to a reviewer. I allow the application's proximity to be unknown to the committee and show that all the same qualitative features of this model continue to hold. See the end of this section for a proof. Neither the committee nor the reviewer observes Q^*, but the reviewer observes a signal Q_P about Q^*. I assume that a related reviewer has greater expertise, meaning that Q_1 gives a more precise signal than Q_0.[24]

After observing the signal, the reviewer sends a message to the committee about the application's quality and the committee then decides whether to approve the grant. When determining what message to send, a reviewer considers his payoffs: for an unrelated application, this is identical to that of the committee, but for a related application, the reviewer now receives $Q^* + B$ if the application is funded and U otherwise. The term B represents his bias. The timing is as follows:

1) An application with true quality Q^* is assigned to a reviewer.
2) The application's type ($P = 1$ or $P = 0$) is determined and is publicly observed.
3) The reviewer observes the signal Q_P.
4) The reviewer sends a costless and unverifiable message M to the committee from some message space $\mathbf{M}$.
5) The committee, observing M, makes a decision $D \in \{0, 1\}$ of whether to fund the grant.
6) True quality is revealed and the reviewer and committee both receive their payoffs.

[24]For simplicity, I assume that the signals Q_P are real numbers with continuous unconditional distributions such that $E(Q^*|Q_P)$ is increasing in Q_P.

Proposition 1 describes the perfect Bayesian equilibria of this game.[25]

PROPOSITION 1: *The equilibria of the game is summarized by the following two cases:*
CASE 1*: $P = 0$. There exists a unique informative equilibrium in which*

1) *The reviewer reports a message Y if $E(Q^*|Q_0) > U$ and N otherwise.*[26]
2) *The committee funds the grant if and only if the message is Y.*

CASE 2*: $P = 1$. There exists a level of bias $B^* > 0$ such that for bias $B \leq B^*$ there is a unique informative equilibrium such that*

1) *The reviewer reports a message Y if $E(Q^*|Q_1) > U - B$ and N otherwise.*
2) *The committee funds the grant if and only if the message is Y.*

When $B > B^*$, only uninformative equilibria exist and the grant is never funded.

PROOF:
Proofs are included at the end of this section.

Proposition 1 says that when bias is sufficiently small, review committees are willing to take the advice of the reviewer because they value her expertise, in spite of the her bias. The committee's decision rule in the informative equilibria of this model is given by

$$D = \underbrace{\mathbb{I}(E(Q^*|Q_0) > U)}_{\text{baseline for unrelated}} + \underbrace{[\mathbb{I}(U > E(Q^*|Q_1) > U - B)]}_{\text{bias for proximate (+)}} P + \underbrace{[\mathbb{I}(E(Q^*|Q_1) > U) - \mathbb{I}(E(Q^*|Q_0) > U)]}_{\text{additional information for proximate (+/-)}} P. \tag{F1}$$

The first term of Equation (F1) indicates that committees listen to advice about unrelated applications. The middle term represents the impact of bias on funding decisions. In particular, lower quality applications (those with $U > E(Q^*|Q_1) > U - B$) will be funded if the applicant is related. The final term represents the impact of information. $\mathbb{I}(E(Q^*|Q_1) > U)$ is the decision that an unbiased reviewer would make, given the lower variance signal of the proximate reviewer. $\mathbb{I}(E(Q^*|Q_0) > U)$ is the decision she actually makes; the difference represents the change in funding outcomes that is due only to better information. Bias decreases the expected quality of funded applications while expertise increases it. The net effect of proximity on the quality of decisions is thus ambiguous.

Equation (F1) demonstrates why differences in funding likelihood among applicants with the same quality need not be due to bias. In particular, the difference in the expected

[25] There are always uninformative equilibria in which messages are meaningless and the grant is never funded. This proposition therefore focuses on informative equilibria, i.e. those in which the committee's decision depends on the reviewer's message. An informative equilibrium is unique if all other informative equilibria are payoff-equivalent for the parties.

[26] I assume there are at least two elements in the message space $\mathbf{M}$ which, without loss, I call Y and N.

likelihood of funding between related and unrelated applications of the same quality is given by

$$\begin{aligned} E[D|Q^*, P=1] - E[D|Q^*, P=0] &= \Pr(U > E(Q^*|Q_1) > U - B) \\ &\quad + \Pr(E(Q^*|Q_1) > U) - \Pr(E(Q^*|Q_0) > U). \end{aligned}$$

This expression will be non will be nonzero even if reviewers are unbiased ($B = 0$). This is because reviewers can more confidently attest to the quality of intellectually related applications, meaning that committees update more following a favorable review. Distinguishing between bias and information driven explanations is important because they have different implications for whether proximity enhances the quality of peer review.

Proof of Proposition 1. — A perfect Bayesian equilibrium for this game is characterized by a message strategy for the reviewer, a set of beliefs about Q^* by the committee for each message, and a decision strategy for the committee. Having defined the equilibrium concept, I proceed with the proof of Proposition 1.

Case 1. Suppose that the reviewer reports her exact posterior and the committee to believes it. In this case, the committee maximizes its utility by funding the proposal if and only if $Q_0 > U$. The reviewer has no incentive to deviate from this strategy because she is receiving her highest payoff as well.

Suppose, now, that there were another informative equilibrium. Each message $M \in \mathbf{M}$ induces a probability of funding $D(M)$. Let the messages be ordered such that $D(\mathbf{M}_1) \leq \cdots \leq D(\mathbf{M}_K)$ where $\mathbf{M}_i$ are the set of messages M_i that induce the same probability of funding $D(M_i)$. For reviewers of type $E(Q^*|Q_0) > U$, the reviewer strictly prefers that the grant be funded. She thus finds it optimal to send the message $\mathbf{M}_K$ that maximizes the probability that the grant is funded. Call this set Y. For $E(Q^*|Q^* + \varepsilon_0) < U$ the reviewer strictly prefer $E(Q^*|Q_0) = U$. Because the distribution of Q_P is assumed to be continuous on $\mathbb{R}$ and such that $E(Q^*|Q_P)$ is increasing in Q_P, this occurs with probability zero. Thus, with probability one, the space of possible messages is equivalent to $\mathbf{M} = \{Y, N\}$. For this equilibrium to be informative, it must be that $D(N) < D(Y)$. Given this, the committee's optimal reaction is to fund when $M = Y$ and to reject otherwise.

If the we allow uninformative equilibria, $D(\mathbf{M}_1) = \cdots = D(\mathbf{M}_K)$ and any reviewer message is permissible. It must be that $D(M_i) = 0$ for all M_i because the outside option U is assumed to be greater than the committee's prior on quality.

Case 2. Now consider the case of a reviewer evaluating a related application. As in Case 1, the set of messages is equivalent, with probability one, to $\mathbf{M} = \{Y, N\}$. In this case, however, reviewers of type $E(Q^*|Q_1) > U - B$ send $M = Y$ and reviewers of type $E(Q^*|Q_1) < U - B$ send $M = N$. The only reviewer who sends any other message is one for which $E(Q^*|Q_1) = U - B$.

Given this messaging strategy, a committee's expectation of Q^* given $M = N$ is $E(Q^*|E(Q^*|Q_1) <$

$U-B$). Since this is less than U, the grant goes unfunded. The committee's expectation of Q^* given $M = Y$ is $E(Q^*|E(Q^*|Q_1) < U - B)$. When this is larger than U, the committee listens to the reviewer's recommendation and we can verify that $D(Y) > D(N)$. When $E(Q^*|E(Q^*|Q^* + \varepsilon_1) < U - B) < U$, the grant is never funded: $D(Y) = D(N) = 0$. In this case, only babbling equilibria exist.

If the we allow uninformative equilibria, $D(\mathbf{M}_1) = \cdots = D(\mathbf{M}_K)$ and any reviewer message is permissible. It must be that $D(M_i) = 0$ for all M_i because the outside option U is assumed to be greater than the committee's prior on quality.

Unobserved proximity: Next, I consider a modification of Proposition 1 where the committee cannot observe whether the application is related to the reviewer.

PROPOSITION A.2: *Assume that p is the probability that an application is related to a reviewer. Then, for every p, there exists a level of bias, B^*, such that for $B < B^*$ there is a unique informative equilibrium:*

The reviewer reports a message Y if his posterior, $E(Q^*|Q_1)$, is greater than $U - B$ and N otherwise.

1) *An unrelated reviewer reports a message Y if his posterior, $E(Q^*|Q_0)$, is greater than U and N otherwise.*
2) *A related reviewer reports a message Y if his posterior, $E(Q^*|Q_1)$, is greater than $U - B$ and N otherwise.*
3) *The committee funds the grant if and only if the message is Y.*

For $B \geq B^*$, only uninformative equilibria exist and the grant is never funded.[27]

PROOF:

In this case, the reviewer's messaging strategy remains the same as in Proposition 1: because reviewers themselves know whether they are proximate, they form, with probability one, strict preferences about whether an application should be funded. Proximate reviewers for which $E(Q^*|Q_1) > U - B$ send $M = Y$ and those for which $E(Q^*|Q_1) < U - B$ send $M = N$. Similarly, unrelated reviewers of type $E(Q^*|Q_0) > U$ send $M = Y$ and unrelated reviewers of type $E(Q^*|Q_0) < U$ send $M = N$.

The committee, however, does not observe the proximity and, as such, forms the following expectation of quality conditional on observing $M = Y$:

$$K\left[E(Q^*|E(Q^*|Q_0) > U)\right] + (1-K)\left[E(Q^*|E(Q^*|Q_1) > U - B)\right]$$

The first term $E(Q^*|E(Q^*|Q_0) > U)$ is the committee's expectation of quality if it knows that the $M = Y$ message is sent by an unrelated reviewer. Similarly, the second

[27] Again, in all cases where an informative equilibrium exists, there also exist uninformative equilibria where the grant is never funded.

term $E(Q^*|E(Q^*|Q_1) > U - B)$ is the committee's expectation of quality if it knows that the message is sent by a related reviewer. The term K is the probability that the committee believes a Y message comes from an unrelated reviewer, that is, $K = E(P = 0|M = Y)$. By Bayes' Rule, this is given by $K = E(P = 0|M = Y) = \frac{E(P=0,M=Y)}{E(M=Y)}$. The overall probability of a Y message is thus given by

$$E(M = Y) = (1 - p)(E(Q^*|Q_0) > U) + p(E(Q^*|Q_1) > U - B)$$

Similarly, the probability that the message is Y and the reviewer is unrelated is given by $(1 - p)(E(Q^*|Q_0) > U)$. As such, we have

$$K = \frac{(1 - p)(E(Q^*|Q_0) > U)}{(1 - p)(E(Q^*|Q_0) > U) + p(E(Q^*|Q_1) > U - B)}.$$

and for

$$K\left[E(Q^*|E(Q^*|Q^* + \varepsilon_0) > U)\right] + (1 - K)\left[E(Q^*|E(Q^*|Q^* + \varepsilon_1) > U - B)\right] > U$$

the committee funds the application. Again, we can verify that $D(Y) > D(N)$. For any fixed p, the threshold B^* can be defined to set this expression equality. There also exist uninformative equilibria where all grants are rejected. This term is less than U, then the grant is never funded: $D(Y) = D(N) = 0$. In this case, only babbling equilibria exist.

F2. *Statistical framework*

The decision rule described by Equation (F1) in the theoretical model can be thought of as a data generating process for the funding decisions I observe. To make this more tractable, I make the following simplifying assumptions: for $P = 0, 1$, the reviewer's signal Q_P can be written as $Q_P = Q^* + \varepsilon_P$ where $\varepsilon_P \sim U[-a_P, a_P]$ and $E(Q^*|Q_P)$ can be approximated by λQ_P for some constant λ_R. Given this, an application's conditional likelihood of funding can be expressed as:

$$\begin{aligned}
E[D|Q^*, P] &= \Pr(\lambda_0(Q^* + \varepsilon_0) > U) + \Pr(U > \lambda_1(Q^* + \varepsilon_1) > U - B)P \\
&\quad + \left[\Pr(\lambda_1(Q^* + \varepsilon_1) > U) - \Pr(\lambda_0(Q^* + \varepsilon_0) > U)\right] P \\
&= \frac{a_0 - U/\lambda_0 + Q^*}{2a_0} + \frac{B}{2a_1\lambda_1}P + \left[\frac{a_1 - U/\lambda_1 + Q^*}{2a_1} - \frac{a_0 - U/\lambda_0 + Q^*}{2a_0}\right] P \\
&= \frac{1}{2} + \underbrace{\frac{1}{2a_0}}_{\text{Quality corr.}} Q^* + \underbrace{\frac{B}{2a_1\lambda_1}}_{\text{Bias term}} P + \underbrace{\left[\frac{1}{2a_1} - \frac{1}{2a_0}\right]}_{\text{Add. corr. for proximate}} PQ^* \\
&\quad - \frac{U}{2a_0\lambda_0} + \left[\frac{1}{2a_0\lambda_0} - \frac{1}{2a_1\lambda_1}\right] PU. \qquad \text{(F2)}
\end{aligned}$$

This distributional assumption has the benefit of allowing me to express the value of bias and expertise in a simple linear way: bias enters the level effect of relatedness while expertise enters the interaction effect. However, the assumption itself is restrictive: having a limited support of the error distribution means that if an application is extremely high (low) quality, the committee will choose to approve (reject) it regardless of what the reviewer says. As such, Equation (F2) is valid for candidates with quality such that $Q^* + \varepsilon_P$ cannot be greater than U or less than U for all possible ε_P. Effectively, this restricts our analysis to grants that are at the margin of funding.

Given these caveats, Equation (F2) shows how I separately identify the role of bias and expertise. In particular, consider the regression analogue of Equation (F2):

$$D = \alpha_0 + \alpha_1 Q^* + \alpha_2 P + \alpha_3 PQ^* + \alpha_4 U + \alpha_5 PU + X\beta + \epsilon, \tag{F3}$$

where X includes other observable I can condition on.

Here, α_2, the coefficient on proximity P, tests for bias: it is nonzero if and only if $B \neq 0$, where B is the bias parameter from the model. Second, the coefficient on PQ^* tests for expertise. To see this, notice that α_1 captures, for unrelated applicants, how responsive funding decisions are to increases in quality. In the model, this is determined by the precision of the reviewer's signal of quality for unrelated applications. The coefficient on PQ^*, meanwhile, captures the additional correlation between quality and funding for related applicants. A high coefficient on PQ means that a committee is more sensitive to increases in the quality of related applicants than to increases in the quality of unrelated applicants. In the model, this is determined by the difference in the precision of signals for related and unrelated applications.

The intuition for separately identifying bias and expertise is the following: if I find that related applications are more (or less) likely to be funded regardless of their quality, then this is a level effect of proximity that I attribute to bias in the NIH funding process. If I find that quality is more predictive of funding among related rather than unrelated applicants, then I conclude that study sections have better information about proposals from related applicants. I do not make any assumptions about the presence, extent, or direction of any potential biases nor do I assume that reviewers necessarily have better information about related applications. Rather, this statistical framework is designed to estimate this.[28]

Finally, the terms U and PU control for funding selectivity; for high cutoffs U, the correlation between funding and quality will be low even in the absence of bias or differential information because the marginal unfunded application is already very high-quality. The RU term, meanwhile, ensures that relationships are not credited for changing the correlation between funding and quality simply by lowering the threshold at which grants are funded.

Equation (F2) says that, as long as Q^* is perfectly observed, exogenous variation in

[28]These predictions hold when reviewers and committees are in an informative equilibrium. If the equilibrium were not informative, then advice from related reviewers would not be taken; I would find no effect of bias and a lower correlation between funding and quality for related applications. My results are not consistent with a non-informative equilibrium.

proximity is not needed to identify the presence of bias. This is because exogenous variation in proximity is necessarily only when aspects of an application's quality are potentially omitted; if quality were observed, one could directly control for any correlation between proximity and quality.

In practice, however, I do not observe an application's true quality Q^*. Instead, I observe a noisy signal $Q = Q^* + v$. Thus, instead of estimating Equation (F3), I estimate

$$D = a_0 + a_1 Q + a_2 R + a_3 RQ + a_4 U + a_5 RU + Xb + e. \tag{F4}$$

Measurement error in quality can potentially pose problems for identification. Proposition 2 describes the conditions that must be met in order to consistently estimate bias from observed data.

PROPOSITION 2: *Given observed quality $Q = Q^* + v$, the bias parameter α_2 in Equation* (F3) *is consistently estimated by a_2 in Equation* (F4) *when the following conditions are met:*

1) *$Cov(P, Q^*|U, PU, X) = 0$ and $Cov(P^2, Q^*|U, PU, X) = 0$,*
2) *$E(v|U, PU, X) = 0$,*
3) *$Cov(v, P|U, PU, X) = 0$.*

PROOF:
:

Condition 1 requires that my measure of proximity, P, be uncorrelated, conditional on observables, with true application quality. If this were not the case, any mismeasurement in true quality Q^* would bias estimates of α_2 through the correlation between Q^* and P. Thus, in my study, exogenous variation in proximity is required only to deal with measurement error.

Condition 2 requires that measurement error be conditionally mean zero. This means that, after controlling for observable traits of the application or applicant, my quality measure cannot be systematically different from what committees themselves are trying to maximize. Otherwise, I may mistakenly conclude that committees are biased when they are actually prioritizing something I do not observe but which is not mean zero different from my quality measure.

Finally, Condition 3 requires that the extent of measurement error not depend, conditional on observables, on whether an applicant is related to a reviewer. This may not be satisfied if related applicants are more likely to be funded and funding itself affects my measure of quality.

PROOF OF PROPOSITION 2. — Measurement error in Q^* can potentially affect the estimation of α_2 in Equation (F3). The presence of U, PU, and X, however, will not affect consistency; for simplicity, I rewrite both the regression suggested by the model and the actual estimating equation with these variables partialed out. The remaining variables should then be thought of as conditional on U, PU, and X

$$D = \alpha_0 + \alpha_1 Q^* + \alpha_2 P + \alpha_3 PQ^* + \epsilon$$

$$\begin{aligned} D &= a_0 + a_1 Q + a_2 P + a_3 PQ + e \\ &= a_0 + W + a_2 P + e, W = a_1 Q + a_3 PQ \end{aligned}$$

The coefficient a_2 is given by:

$$a_2 = \frac{\mathrm{Var}(W)\mathrm{Cov}(D,P) - \mathrm{Cov}(W,P)\mathrm{Cov}(D,W)}{\mathrm{Var}(W)\mathrm{Var}(P) - \mathrm{Cov}(W,P)^2}$$

Consider $\mathrm{Cov}(W,P)$:

$$\begin{aligned} \mathrm{Cov}(W,P) &= \mathrm{Cov}(a_1(Q^* + v) + a_3 P(Q^* + v), P) \\ &= a_1 \mathrm{Cov}(Q^*, P) + a_1 \mathrm{Cov}(v, P) + a_3 \mathrm{Cov}(PQ^*, P) + a_3 \mathrm{Cov}(Pv, P) \end{aligned}$$

Under the assumption that P and Q^* are conditionally independent, this yields:

$$\begin{aligned} \mathrm{Cov}(W,P) &= a_3 \mathrm{Cov}(PQ^*, P) + a_3 \mathrm{Cov}(Pv, P) \\ &= a_3 \left[E(P^2 Q^*) - E(PQ^*)E(P)\right] + a_3 \left[E(P^2 v) - E(Pv)E(P)\right] \\ &= a_3 \left[E(P^2)E(Q^*) - E(P)^2 E(Q^*)\right] + a_3 \left[E(P^2)E(v) - E(P)^2 E(v)\right] \\ &= a_3 \left[E(P^2)0 - E(P)^2 0\right] + a_3 \left[E(P^2)0 - E(P)^2 0\right] \\ &= 0 \end{aligned}$$

With this simplification, the expression for the estimated coefficient on a_2 becomes:

$$\begin{aligned} a_2 &= \frac{\mathrm{Var}(W)\mathrm{Cov}(D,P) - \mathrm{Cov}(W,P)\mathrm{Cov}(D,W)}{\mathrm{Var}(W)\mathrm{Var}(P) - \mathrm{Cov}(W,P)^2} \\ &= \frac{\mathrm{Var}(W)\mathrm{Cov}(D,P)}{\mathrm{Var}(W)\mathrm{Var}(P)} \\ &= \frac{\mathrm{Cov}(D,P)}{\mathrm{Var}(P)} \\ &= \frac{\mathrm{Cov}(\alpha_0 + \alpha_1 Q^* + \alpha_2 P + \alpha_3 PQ^* + \varepsilon, P)}{\mathrm{Var}(P)} \\ &= \frac{\alpha_2 \mathrm{Var}(P) + \alpha_3 \mathrm{Cov}(PQ^*, P)}{\mathrm{Var}(P)} \\ &= \frac{\alpha_2 \mathrm{Var}(P) + \alpha_3 \left[E(P^2)E(Q^*) - E(P)^2 E(Q^*)\right]}{\mathrm{Var}(P)} \\ &= \alpha_2 \end{aligned}$$

F3. Empirical Estimates

Appendix Table O estimates the regression equation suggested by Equation (F4). Specifically,

$$
\begin{aligned}
Assessment_{icmt} \quad = \quad & a_0 + a_1 \text{Proximity to Permanent}_{icmt} \\
& + a_2 \text{Proximate to Permanent}_{icmt} \times \text{Quality}_{icmt} \\
& + a_3 \text{Quality}_{icmt} + a_4 \text{Total Proximity}_{icmt} \\
& + a_5 \text{Total Proximity}_{icmt} \times \text{Quality}_{icmt} \\
& + \mu X_{icmt} + \delta_{cmt} + \varepsilon_{icmt}.
\end{aligned} \tag{F5}
$$

I am interested in the coefficients a_1 and a_2. Proximity to Permanent$_{icmt}$ is defined as the number of permanent reviewers that cite an applicant's prior work. a_1 captures the effect of proximity on funding that is attributable to bias: does being cited by permanent reviewers, conditional on total proximity, affect an applicant's likelihood of being funded for reasons unrelated to quality? Bias is identified as the change in the *level* probability that a proximate applicant is funded. Meanwhile, Proximate to Permanent$_{icmt}$ × Quality$_{icmt}$ is the interaction of an application's quality with an indicator for whether an applicant has been cited by a permanent reviewer. The coefficient a_2 captures the role of expertise: it asks whether there is a steeper *slope* in the relationship between quality and funding for applicants with intellectual ties to more influential reviewers.

The remaining variables in Equation (F5) control for potentially contaminating variation. I control for the level of effect of application quality, total proximity to all reviewers, as well as the interaction between these two terms. Controlling for these terms means that the coefficient of interest a_1 and a_2 are estimated from applicants who have been cited by the same total number of reviewers, but who differ in their ties to permanent reviewers. I also control for a variety of past publication and demographic characteristics, X_{icmt}, described in Section III.

Finally, the model in Appendix F that motivates Equation (F5) also requires that I include controls for the degree of selectivity in a committee. When committees a very small percentage of applicants, the correlation between funding and quality will be low even in the absence of bias or differential information because the marginal unfunded application is already very high-quality. In my empirical implementation, I proxy for selectivity using the percentile pay line of the committee and include a level control for pay line (this is absorbed in the meeting fixed effect). I also control for the interaction of proximity and the payline. This ensures that proximity is not credited for changing the correlation between funding and quality simply by lowering the threshold at which grants are funded. My results are not affected by either the inclusion or exclusion of these variables.

Appendix Table O reports my estimates of Equation (F5), decomposing the effects of bias and expertise. Column 2 reports estimates of the coefficients from Equation (F5) for funding status. The positive and significant coefficients on the level effect of proximity

(0.0068) indicates that reviewers are biased in favor of applicants and the positive and significant coefficients on the interaction of proximity with quality (0.076) indicate that reviewers also have more expertise about related applications. Reviewers, however, also do a better job of discerning quality of related applicants. Consider a 1 standard deviation (51 citations) increase in the quality of a grant application: for an applicant cited by a single permanent reviewer, my estimates imply that this change would increase her chances of funding by $(0.0136 + 0.0176 - 0.0005) * 0.51 * 100 = 1.6$ percentage points or 1.6/21.4=7.5 percent. If, instead, this applicant has been cited by a single temporary reviewer, the same increase in quality would only increase her chances of funding by $(0.0136 - 0.0005) * 0.51 * 100 = 0.7$ percentage points or 3.3 percent. Committees are twice as responsive to changes in the quality of applications in the subject area of permanent members.

Appendix Table O: What is the effect of proximity? Linear Specification

	1(Score is above the payline)		**Score**		**1(Scored at all)**	
	Mean = 0.214, SD = 0.410		Mean = 71.18, SD = 18.75		Mean = 0.640, SD = 0.480	
	(1)	(2)	(3)	(4)	(5)	(6)
Proximity to Permanent Reviewers	0.0072***	0.0068***	0.2736***	0.2590***	0.0047**	0.0043**
	(0.002)	*(0.002)*	*(0.094)*	*(0.095)*	*(0.002)*	*(0.002)*
Proximate to Permanent Reviewers × Grant Application Quality		0.0176**		0.2739		0.0162*
		(0.008)		*(0.325)*		*(0.009)*
Grant Application Quality		0.0136**		0.5568**		0.0305***
		(0.006)		*(0.261)*		*(0.008)*
Total Proximity X Grant Application Quality		-0.0005		-0.0043		-0.0036***
		(0.001)		*(0.049)*		*(0.001)*
Observations	93,558	93,558	57,613	57,613	93,558	93,558
R-squared	0.0935	0.0949	0.1426	0.1431	0.1312	0.1322
Meeting FEs	X	X	X	X	X	X
# of Proximate Reviewer FEs	X	X	X	X	X	X
Past Performance, Past Grants, and Demographics	X	X	X	X	X	X

Notes: See notes to Table 1 for details about the sample. Coefficients are reported from a regression of committee decisions (above payline, score, or scored at all) on relatedness and quality measures, controlling for meeting level fixed effects. Proximity to permanent reviewers is defined as the number of permanent reviewers who have cited the applicant's research in the 5 years prior to grant review. "Grant Application Quality" is defined as the number of citations up to 2008, for all publications that are text-matched to the grant application within 1 year of grant review, in the 100s unit. "Past Performance, Past Grants, and Demographics" include indicators for sex and whether an applicant's name is Hispanic, East Asian, or South Asian, indicator variables for deciles of an applicant's total number of citations and publications over the past 5 years, indicators for whether an applicant has an M.D. and/or a Ph.D., and indicators for the number of past R01 and other NIH grants an applicant has won, as well as indicators for how many she has applied to.

*** p<0.01, ** p<0.05, * p<0.1

F4. Efficiency

Finally, Equation (F5) allows me to construct counterfactual funding decisions, made in the absence of proximity that would have been obtained in the absence of relationships. Specifically, I define

$$\begin{aligned} \text{Funding}_{icmt}^{\text{Benchmark}} &= \text{Funding}_{icmt} \text{ (actual funding)} \\ \text{Funding}_{icmt}^{\text{No Proximity}} &= \text{Funding}_{icmt} - \widehat{a_1}\text{Total Proximate Permanent}_{icmt} \\ &- \widehat{a_2}\text{Quality}_{icmt} \times \text{Proximate to Permanent}_{icmt}, \end{aligned}$$

where $\widehat{a_1}$ and $\widehat{a_2}$ are estimated from Equation (F5).[29] The counterfactual funding decision represents what the committee would have chosen had applicants related to permanent members been treated as if they were unrelated.

I summarize the effect of relationships by comparing the quality of the proposals that would have been funded had relationships not been taken into account with the quality of those that actually are funded. Specifically, I consider all applications that are funded and sum up the number of publications and citations that accrue to this portfolio. This is my benchmark measure of the quality of NIH peer review. I then simulate what applications would have been funded had relationships not been taken into account. To do this, I fix the total number of proposals that are funded in each committee meeting but reorder applications by their counterfactual funding probabilities. I sum up the number of publications and citations that accrue to this new portfolio of funded grants. The difference in the quality of the benchmark and counterfactual portfolio provides a concrete, summary measure of the effect of relationships on the quality of research that the NIH supports.

Appendix Table P estimates the effect of relationships on the quality of research that the NIH supports. In effect, I ask what the NIH portfolio of funded grants would have been had committees treated applicants who are related to permanent members as if they were not, holding all else fixed. In my sample, I observe 93,558 applications, 24,404 of which are funded. Using this strategy, I find that 2,500, or 2.7 percent, of these applications change funding status under the counterfactual.

On average, working in the same area as influential reviewers helps an applicant obtain funding; ignoring this intellectual connection would decrease the number of proximate applicants who are funded by 3.0 percent. The quality of applications funded when intellectual proximity is taken into account, however, is higher. The overall portfolio of funded grants under the counterfactual produces two to three percent fewer citations, publications, and high-impact publications. To take account of the fact that some grants are funded and others are not, I use my standard funding-purged measure of grant application quality—text-matched publications within one year of grant review, and citations to those publications—as the measure of grant output used for this analysis. This has the benefit

[29] Even though $\text{Funding}_{icmt}^{\text{No Relationship}}$ is constructed using estimates from Equation (F5), it does not rely on the model to interpret those coefficients.

of allowing me to compare the benchmark NIH portfolio with counterfactual results, holding constant the effect of actual funding status. However, a downside of this approach is that the stringent matching requirement will undercount the total number of publications (and therefore citations) associated with these grants. This exercise should thus be used to compare the percentage difference between the benchmark and counterfactual no-proximity cases, rather than to discern the level of NIH output.

Appendix Table P: What is the effect of proximity on the aggregate quality of NIH funded grants?

	Benchmark	**No Proximity**
Number of Funded Grants	24,404	24,404
Number of Grants that Change Funding Status	2,500	2,500
Total # Citations	584,124	566,284
(% change relative to benchmark)		*(3.05)*
Total # Publications	11,149	10,851
(% change relative to benchmark)		*(2.67)*
Total # in Top 99% of Citations	590	572
(% change relative to benchmark)		*(3.05)*
Total # in Top 90% of Citations	10,239	9,925
(% change relative to benchmark)		*(3.07)*
Total # Related Applicants Funded	18,666	18,113
(% change relative to benchmark)		*(2.96)*

Notes: Benchmark refers to characteristics of grants ordered according to their predicted probability of funding, using the main regression of funding status on proximity and grant application quality. ″Benchmark″ figures are the grant quality measures for a grants that would be funded if we used the predicted ordering from the regression of funding likelihood on relatedness and quality estimated in Appendix Table L. ″No proximity″ refers to the predicted ordering of grants under the same regression, but under the assumption that relatedness to permanent members and relatedness to permanent members interacted with quality do not matter (their coefficients are set to zero). To take account of the fact that some grants are funded and others are not, we use our standard funding-purged measure of grant application quality: text-matched publications within one year of grant review, and citations to those publications. The number of projects that are funded is kept constant within meeting. See text for details.