Growing Stem Cells:
The Impact of Federal Funding Policy on the U.S. Scientific Frontier*

Jeffrey L. Furman
Boston University & NBER

Fiona Murray
MIT Sloan School of Management

Scott Stern
MIT Sloan School of Management & NBER

January 17, 2012

Journal of Policy Analysis & Management
(forthcoming)

* Acknowledgements: We thank David Altshuler, Pierre Azoulay, Kevin Eggan, Amy Finkelstein, Joshua Gans, Julia Lane, Chris Liu, Megan MacGarvie, Joseph Newhouse, Matt Notowidigdo, Alex Oettl, Bhaven Sampat, Eunhee Sohn, Paula Stephan, Jerry Thursby, Marie Thursby, and participants in numerous conferences and seminars for comments and suggestions. Lisa Bassett, Devin Fensterheim, Pegi Fogertey, Anna Harrington, Oren Klachkin, Josh Krieger, Brianna Pesci, and Alexandra Somers provided excellent research assistance. All errors are our own. Financial support for this research was provided by the National Science Foundation, under grant #0738394. Author contact information: Jeffrey L. Furman, Boston University School of Management, 595 Commonwealth Ave – #653a, Boston, MA 02215, furman@bu.edu; Fiona Murray, MIT Sloan School of Management, 50 Memorial Drive, E52-568, Cambridge, MA 02142, fmurray@mit.edu and Scott Stern, MIT Sloan School of Management, 50 Memorial Drive, E52-568, Cambridge, MA 02142, sstern@mit.edu.

Growing Stem Cells:
The Impact of Federal Funding Policy on the U.S. Scientific Frontier*

Abstract

This paper articulates a citation-based approach to science policy evaluation and employs that approach to investigate the impact of the United States' 2001 policy regarding the Federal funding of human embryonic stem cell (hESC) research. We evaluate the impact of the policy in the level of U.S. hESC research, the U.S. position at the knowledge frontier, and the strategic response of U.S. scientists. Consistent with recent research on the science of science and innovation policy, we employ a difference-in-differences approach using bibliometric data with the aim of analyzing the *causal* impact of the policy on cumulative research. Our estimates suggest that in the aftermath of the 2001 policy, U.S. production of hESC research lagged 35-40 percent behind anticipated levels. However, this relative decline was largely concentrated in the years 2001-2003 and ameliorated over time. The rebound in U.S. hESC research was driven by contributions by researchers at elite U.S. institutions and U.S. researchers who collaborated with international partners. The results suggest that scientists respond strategically to research funding restrictions, and that modest science policy shifts can have a significant influence on the within-country composition of research and the pattern of global research collaboration.

Key Words: stem cells, science of science and innovation policy, scientific productivity, R&D management, innovation

I. Introduction

This paper evaluates whether conditions placed on U.S. Federal funding of human embryonic stem cell (hESC) research, as reflected in the August 2001 policy decision of the Bush Administration, influenced the level of U.S. hESC scientific activity and the U.S. position at the knowledge frontier. We assess the strategic response of U.S. scientists to this policy intervention, specifically examining changes in levels of U.S. research as well as patterns of international collaboration and research output by high and low status universities.

Taking advantage of recent advances in the use of difference-in-differences methods for science policy evaluation, our analysis employs a citation-based approach to establish the causal impact of the 2001 hESC policy decision. Our paper thus makes three contributions. First, we provide direct evidence regarding a key policy question – what was the impact of the 2001 stem cell policy decision on the evolution of hESC research in the United States and abroad? Second, we address a more general question of both theoretical and policy interest – how does the targeted funding of research influence the direction and evolution of scientific research? Third, our analysis offers a primer regarding the opportunities and challenges associated with a citation-based difference-in-differences approach to science policy analysis.

Our analysis is motivated by a central question in science policy: How do conditions on scientific funding matter for scientific progress? On the one hand, a long tradition in science policy emphasizes that the evolution of particular research fields or disciplines is guided by the internal logic of cumulative discovery and the social structure and norms of scientific communities (Kuhn, 1962; Merton, 1973; David and Dasgupta, 1994). At the same time, funders – from governments to foundations to corporations – have long attempted to shape the direction of scientific progress by conditioning funding on particular research approaches or directions (Dasgutpa and David, 1994; Fuchs, 2009, 2010; and see Gans and Murray 2011 for a recent review). While the aggregate long-term level of funding for a broad field of course matters (Stephan, 1996; Adams and Griliches, 1996), there is much less evidence as to the impact of specific conditions attached to funding on the evolution of a scientific field. Whereas many limitations on scientific funding are developed in conjunction with the scientific community (e.g., rules regarding informed consent), other policies are imposed on the scientific community based on the politics or preferences of funders. In this latter case, the intended impact of targeted funding may be mitigated to a significant extent through strategic behavior on the part of scientists (Aghion, Dewatripont, and Stein, 2008; Murray, 2010; Gans and Murray, 2011). The possibility of strategic behavior undermining the intent of targeted funding (or restrictions on funding) is likely to be most salient in settings where the focal scientific agenda is considered to be "hot" by the scientific community, when alternative sources of funding

may be available, and when scientists largely disagree with the rationale for the funding policy. The key objective of this paper is to evaluate the impact of science policy funding conditions in precisely such a circumstance.

We specifically examine the impact of the 2001 Bush Administration hESC policy on cumulative lines of research that built on the hESC research breakthroughs taking place from 1998 through 2001. This is not the first project to examine the impact of hESC policy (see Owen-Smith and McCormick, 2006; Levine, 2005, among others). However, our approach is distinct from prior work in developing a citation-based science policy evaluation approach that identifies the precise *causal* linkage between the U.S. policy and the evolution of subsequent scientific research. Our approach builds on citation-based evaluative tools for the "Science of Science and Innovation Policy" (Murray and Stern, 2007; Furman and Stern 2010; Murray et al., 2011).

Our citation-based approach combines three distinct elements. First, rather than using the count of publications to trace out activities in hESC research in general, we develop a dataset of *core* hESC publications based on a report produced by the National Institutes of Health (2001) and then trace out the entire set of *forward citations* to these articles. Second, we develop a comparison group of papers that were unaffected by the hESC policy in order to construct a well-defined group of control articles, allowing us to observe how citations to a similar population of scientific papers evolved over time. Such observations allow us to develop a counterfactual estimate of how the citations to the hESC articles would have evolved absent the policy intervention. While we experiment with several control groups, we focus primarily on a group of seminal research papers in the area of RNA interference (RNAi), a contemporaneous life sciences breakthrough not directly affected by the hESC intervention.[1] By comparing the evolution of hESC citations to those from a control group, we are able to account for factors such as the globalization of science over time or even more specific confounding factors, such as the events of September 11, 2001. Finally, to analyze these data, we apply difference-in-differences econometric techniques to evaluate how the pattern of citations to the hESC articles changed in response to the 2001 policy intervention. Importantly, we do not simply analyze how the policy impacted the level of hESC research, but focus on how the policy changed the geographic distribution of follow-on research, focusing specifically on the U.S. relative to other countries. Additionally, we consider more fine-grained shifts, such as the impact of the policy on researchers in different types of institutions (e.g., high-status versus low-status) or in different types of collaborations (e.g., domestic versus international).

[1] The discovery of RNAi was recognized with the award of the 2006 Nobel Prize in Medicine/Physiology to Andrew Fire and Craig Mello.

Our analysis establishes a number of core findings. First, our results suggest that U.S.-based hESC research experienced a significant decline relative to such research outside the United States in the years immediately following the policy's announcement. This is consistent with prior work, including Owen-Smith and McCormick (2006) and Levine (2008). However, our research also documents a recovery for US-based hESC research after the initial decline in 2001-2003; specifically, the relative output of U.S.-based hESC research recovers between 2004 and 2007. We investigate the channels through which this recovery was achieved. We find that the policy hESC output does not affect researchers at elite U.S. universities in the same way as non-elite universities. Specifically, we find that hESC researchers at these institutions nearly completely recover from the policy shock to hESC output. In addition, we find that collaborations between U.S. authors and researchers outside the United States were impacted to a lesser extent than those papers that did not involve cross-national collaborations. In conjunction with our qualitative research based on discussions with hESC researchers (Furman and Murray, 2008), we interpret these results as implying (a) that researchers at elite U.S. institutions responded to the policy by developing alternate funding mechanisms to support hESC research and (b) that an alternative, and potentially complementary, method to continue contributing to hESC research involved collaboration with researchers in less restrictive policy environments.

The findings in this paper suggest that the scientific community responds strategically to the institutional and policy conditions affecting science (Jaffe, 2006; Murray, 2010). In addition, the paper contributes to emerging research on the science of science policy using policy-analytic tools to identify the impact of an unanticipated policy on the productivity and organization of science (Marburger, 2005; Lane, 2009; Lane and Bertuzzi, 2011). Finally, our results highlight the sensitivity of scientific output to policies which condition research funding in ways that are opposed by the scientific community (Gans and Murray, 2011).

The remainder of the paper is organized as follows: In Section II, we outline the citation-based difference-in-differences approach and compare it to more traditional approaches which focus on tracking research activity over time. Section III introduces the scientific and policy history of human embryonic stem cell research and motivates our specific research questions. Section IV describes the application of the citation-based methodology to the hESC data, and Section V presents our main empirical results. A concluding section considers the implications of our findings for science policy research.

II. The Citation-based Approach to Science Policy Evaluation

II.1. Conceptual foundations of citation-based science policy evaluation

The citation-based approach to policy evaluation focuses on assessing the impact of critical interventions on the process of step-by-step scientific discovery. Our approach builds on a long tradition of the use of bibliometrics in science policy dating back at least to the work of physicist and historian of science Derek de Solla Price (1963, 1965). Traditionally, bibliometric analysis has tended to focus on how different policies shape the overall scientific landscape or channel research efforts into one domain over another. In practice, this type of exercise has been done by counting publications (or patents) associated with particular keywords, topics, journals, institutions, or countries. This approach has been used to assess a wide variety of policy questions, including the role of national institutions on scientific productivity (de Solla Price, 1967), the impact of organizational structures and incentives on scientist activity (Levin and Stephan, 1991), or the interrelationships among patenting and publication by academic scientists (Azoulay, et al., 2007; Ding, Murray and Stuart, 2006).

Our approach builds on this earlier work but leverages the means by which scientists document their processes of step-by-step cumulative scientific discovery and acknowledge the work of those upon whom they build. Specifically, scientific citation and scientific priority are fundamental norms of open science (Garfield, 1955; de Solla Price, 1962; Merton, 1973) that we can use to track the rate and direction of particular scientific research lines. While citations are a highly imperfect (and noisy) tool for tracking the evolution of particular research lines, the act of referencing prior work is both a pervasive scientific norm (with potential professional consequences if not followed) that delimits the contribution of any particular follow-on paper. In other words, given the norms and practices of a particular scientific area, scientists have incentives to cite the direct work that their discoveries build upon but have few incentives to engage in "gratuitous" citation beyond the norms expected by the scientific community. Not simply a measurement tool, the use of a citation-based approach allows us to document process of step-by-step cumulative scientific discovery that has been at the heart of modern treatments of scientific and technical progress in economics and the history of science (among others, Romer, 1990; Scotchmer, 1991; Aghion and Howitt, 1992; Mokyr, 2002; Aghion, Dewatripont and Stein, 2008; Murray, et al., 2011). While citations may sometimes indicate that the prior research is directly being built upon, and in other cases serve to highlight the broad outlines of a field, our approach exploits both types of references to capture how research conducted prior to a policy intervention informs and shapes research activity after the policy intervention. Our objective is to leverage citation data in order to be able to infer the impact of particular policies – such as the 2001 stem cell decision – on the structure and composition of follow-on scientific research citations.

To identify the causal impact of particular policies on the level and structure of follow-on research as measured through citations, we must not simply measure the *change* in the level of citations over time (e.g., before versus after the policy) but must also address what we would have anticipated the citation level and structure to be at any point in time in the absence of the policy shift we seek to evaluate. Thus, the second notable difference between this approach and typical bibliometric approaches is that we employ citation analysis with the aim of identifying the precise *causal* influence of science policy on cumulative research lines. Here we draw on advances in the program evaluation literature (Angrist and Krueger, 2001; Imbens and Wooldridge, 2009) where the aim is to draw causal inference about the impact of specific programs or policy environments by looking at exogenous shocks to those programs for a particular population while also tracking the behavior of a control group of individuals who would be expected to follow a similar evolution but are unaffected by the program change. In other words, the core idea behind our citation-based approach is to evaluate how the evolution of particular research lines *changes* with *changes* in the institutional or policy environment. To do so, we will take advantage of how the level and structure of citations to a group of core articles related to the research line change before and after the policy shift, and we will compare that change to the group of citations to those of control articles that are expected to have a similar evolution over time but are unaffected by the focal policy shift. The remainder of this section outlines this approach in detail.

II.2. Key stages in the application of citation-based evaluation

It is useful to begin our analysis by considering how we would evaluate science policy changes in an ideal experimental setting. In such a setting, a researcher would use citation patterns to identify the causal impact of science policy on scientific progress by assigning knowledge to multiple distinctive policy environments and then compare citation patterns under each policy regime. Almost by construction, however, the ability to conduct classical experiments interrogating the process of scientific discovery is extremely limited given the uniqueness of different scientific trajectories. Absent such controls, assessing the causal influence of policy environments on scientific progress is beset by a fundamental inference problem. Specifically, it is difficult to isolate the impact of a policy (or program) from the (often highly distinctive) characteristics of knowledge that is impacted by the policy. For example, the observation that researchers who receive more long-term unrestricted funding undertake high levels of research or more risky research is hard to interpret in a causal sense. On the one hand it could imply that funding conditions cause distinctive research trajectories (a treatment effect). On the other, this empirical finding might simply reflect the fact that researchers

attracted to unrestricted funds tend to pursue more risky projects (a selection effect) (Azoulay, Graff-Zivin and Manso, 2011).

Our citation-based approach combines four interrelated elements in order to overcome this fundamental inference problem. First, it requires that a policy intervention (or shock) has the potential to affect the future cumulative trajectory of knowledge in specific and well-defined research areas. For example, the retraction of research papers constitutes an intervention (at the journal level not the policy level) that is intended to influence a specific research trajectory (i.e., it aims to inhibit knowledge accumulation based on research acknowledge to be false) (Furman, Jensen, and Murray, 2012). A shock that is ideal from a research standpoint should have a number of characteristics. In order to be worthy of research attention, a shock should either shed light on an important theoretical issue or have a potentially important impact on cumulative research. Another critical element in the effective application of the citation-based approach is that the policy shock be *exogenous*, i.e., unanticipated by the research community it affects: By comparing the patterns of knowledge accumulation before and after an exogenous policy intervention, investigators can more plausibly measure the causal impact of the intervention on scientific progress. If the intervention were anticipated by the research community, then it would be impossible to disentangle with certainty whether research behaviors were changing because of the policy or whether changing research behaviors augur the policy's arrival. It would be ideal if the timing of a shock were unambiguous in its onset, as this would enable science policy evaluation to clearly note the pre- and post-shock periods and, thus, compare research outcomes before and after the shock takes effect. The ideal shock would have an impact that is extremely precise and can be easily observed in follow-on research. If either the timing of the shock or the research community affected by the shock is imprecise, the interpretation of observed results will be murky. For example, if a shock were to apply to a range of research areas rather than a single area, it would be difficult to definitively track the impact of the shock on follow-on research. Finally, for identification purposes, it would also be helpful if the shock were to affect different "units" of knowledge at different times. This is not absolutely necessary, but can be helpful, as this would enable the econometric analysis to separately identify the impact of the shock from the impact of the time period in which it occurred. For example, if a policy is implemented during a particular presidential administration, it may be difficult to disentangle whether the policy intervention or other factors associated with the presidential administration led to the observed outcomes.

In addition to requiring a shock of the type described above, the second element of the citation-based approach involves pairing the treated research area (i.e., the research area hypothesized to be influenced by the policy shift) with a similar research area that is not affected by the policy

intervention and whose pattern of knowledge accumulation can generate a counterfactual estimate of the way in which research would have progressed in the absence of the policy intervention. In order for the control group to generate a counterfactual estimate of the pattern of scientific progress that would have occurred had the policy intervention not occurred a number of criteria must be satisfied: (1) the control group should not be affected by policy shock; (2) it must have features similar to that of treated group prior to shock; and (3) post-shock patterns must be expected to reflect patterns that would have occurred in treatment group in the event that treatment did not occur. In other words, a control group is sought that is similar on observable characteristics but is thought not to be impacted by the policy intervention (treatment). The appropriateness of the control group can be evaluated by comparing bibliometric features; however, it is also essential to have a qualitative understanding of the control group in order to understand the appropriateness of the match. It is particularly helpful to talk to scientists in the "treated" research community to ascertain appropriate control areas of research as well as those in the control community. The unexpected coupling among research areas can be problematic when establishing an appropriate control group.[2]

Following the determination of a treatment and control research areas, the third element in the analysis follows naturally: identifying a set of "core" research articles that define a body of knowledge (in both the treatment and control groups) in the period prior to the policy intervention. For example, to explore knowledge trajectories associated with transgenic mice, we use a comprehensive database called Mouse Genetics Informatics to help match specific research mice to the publications that disclosed information regarding the development and breeding of particular mouse model (Murray et al., 2011). The treatment group in this project included a set of research mice over which intellectual property rights changed during the observation period; the control group included a set of mouse models used similarly beforehand that were not affected by the shift in IP rights.

To measure knowledge accumulation along particular research lines, we track citations to the set of core research papers and interpret forward citations to these articles as evidence of knowledge accumulation along the chosen research line. We acknowledge that citations are noisy measures; however, our interpretation of forward citations in this manner relies on the fact that (a) core

[2] In several papers including this one, we consider a particular type of control group that we refer to as "nearest neighbors' (see Furman and Stern, 2011; Furman, Jensen and Murray, 2011). These papers are the next paper published in the journal (in chronological order) that is not affected by the particular treatment (e.g. a paper that is not associated with research materials or a paper that is not retracted). These neighboring papers can, in some designs, provide a useful match because they meet similar journal quality standards, are of the same age and are generally in the same broad subject area. If the analysis requires an extremely close match in topic space between the treated and control groups, nearest neighbor articles may be a poor match when the sample includes general interest journals, like *Science*, in which articles in physics may occupy space next to articles in very different fields, like biology or social science.

published papers are produced at a specific and measurable point in time and (b) the way they are used by follow-on researchers can be captured and measured in citations with the norms of citation neither explicitly changing over time, nor being shifted by the policy.[3] Here, too, a detailed understanding of the phenomenon, supported by qualitative research, can be extremely helpful. To date, we have found a high fidelity match between core papers and follow-on knowledge accumulation in instances where important research materials are developed and disclosed in the core papers. For example, to trace the way in which knowledge associated with particular life science research materials accumulates, we (a) match individual research materials to the scientific publications that first identify and describe the characteristics of those materials and (b) track forward citations to the those reference articles (Furman and Stern, 2011). By identifying research areas where the cumulative follow-on research trajectory builds on key knowledge disclosed in core publications, and by closely tracing out the citation relationships between the follow-on research and the core publications, we are able to leverage the widespread availability of citation data to evaluate the impact of science policy decisions.

With an exogenous policy intervention identified, mapped to both a treatment and control set of research areas and linked to a set of original reference articles and forward citations (reflecting knowledge accumulation along their research trajectories), causal analysis can proceed. This is typically accomplished using econometric difference-in-differences techniques (although insights can also often be gained from graphical approaches). In this framework, the annual count of forward citations to each "core" research article (treated and controls) is estimated, subject to a variety of year, article age, and article fixed effects. In order to ensure that these are separately identified, it is necessary to have variation in article cohort age, as year and age effects will not be separately identified if all articles are published in the same year. The inability to isolate these effects on follow-on citations will reduce the ability to identify the main effect of interest. It is also ideal, though not always possible, if the policy intervention were implemented at different times for different articles in the affected research line. This is helpful for identifying the impact of the intervention separately from the impact of factors associated with particular points in time, represented econometrically by calendar year fixed effects. For example, the impact of retraction on follow-on research can be identified separately from calendar year fixed effects, as various articles

[3] We recognize that bibliometric analysis provides only a noisy indicator of scientific progress (see, e.g., Garfield, 1979; Lindsey, 1989; and Schubert and Braun, 1993). For example, for a number of reasons, small differences in the citation rate of a single paper (particularly early in its publication history) are of limited value in distinguishing the importance of research or its use by the research community. The citation-based approach to science policy evaluation attempts to minimize the impact of these limitations by drawing comparisons among large samples of publications, comparing across control samples, and assessing the impact of policy changes by drawing comparisons within articles across time.

are retracted in different calendar years. For policies implemented at a specific point in time, however, this identification strategy faces a more severe challenge.

Citation-based science policy evaluation involves the estimation of difference-in-differences models, most often using a count with fixed effects (or conditional fixed effects) for each article, which adjust for the fact that different articles have different propensities to generate follow-on research.[4] The models are often estimated using a negative binomial model because annual citation data is skewed to the right and over-dispersed relative to Poisson. In addition, the rate of citation to a given piece of research will vary with the calendar year and with the time elapsed since initial publication. A key variable in the analysis is the "1/0" policy "shock" variable which effectively switches the policy on and off in any given calendar year. In order to interpret a significant result for this variable as a statement that the policy intervention *caused* changes in the patterns of follow-on research subsequent to the policy intervention, it is important to reiterate that (a) the changes observed following the policy shift did not begin to occur prior to the intervention and (b) the research trajectory considered to be affected by the policy shift (the treatment group) can be usefully compared to a similar research area that is not affected by the policy intervention. This will enable the researcher to identify the "counterfactual," i.e., the estimate of how the treated research line would have developed had the policy intervention not occurred (Imbens and Wooldridge, 2009).

II.3. Citation-based Science Policy Evaluation vs. other approaches

The citation-based approach to science policy evaluation will be possible when each of the elements articulated above is present. This approach offers some advantages (and disadvantages) relative to publication-based (or keyword) approaches more typically used when asking whether/how a policy change (or a change in the institutional environment) impacts the evolution of a stream of cumulative research. It is therefore critical for researchers to determine the most appropriate application of each method.

The approach can be particularly insightful when there is a tight citation-based linkage between specific reference articles and follow-on research. We believe that these linkages are relatively strong as shown in our earlier work on access to biological materials (Furman and Stern, 2011) and research mice (Murray, et al., 2011). With regards to our analysis of the stem cell policy, we make the claim that at least with respect to the core research produced prior to the policy shock, any follow-on research article in this area would, almost by definition, cite at least one (or more) of

[4] When using a conditional fixed effects estimator, one citation year and one age fixed effect are not separately identified (Hall et al., 2007). As long as the main effect of interest is separable from these effects, a specification can be designed to overcome this identification issue.

the reference research articles that were published prior to the policy shock. Indeed, the strong materiality of the hESC field is at once the source of the controversy and an indicator of its suitability for application to the stem cell setting. Of course, even in cases with a weaker link to forward citations, our approach is salient to the extent that both treatment and control groups operate under similar institutional norms for citations. For example, in our analysis of the impact of retractions on knowledge accumulation, both the retracted and non-retracted research are likely to be subject to similar (albeit less powerful) citation norms.

Our methods will be less powerful (and could even be biased) if there is a change to the linkage between follow-on research and a core set of research articles. For example, if the policies of interest actually changed the nature of what was cited (for the reference articles but not for the controls), it is possible that one could misinterpret a change in citation behavior as a substantive shift in research output rather than simply a shift in citation practices. To give one example, one must be careful in evaluating the role of different information technology tools on scholarly publication, as these tools are not simply changing the availability of prior research but also the practices and norms of citation behavior itself. Likewise, if one were to believe that in the area of hESC research, citation practices were strategically changed in the post-policy period to avoid the attention of funders (rather than lower citations being a reflection of lower knowledge accumulation) then the use of citation-based approaches would be less applicable.

A number of additional cautions should be issued in applying the citations-based approach. While the identification of policy shocks and control groups is straightforward in principle, they are in practice often quite difficult. Certainly, the citation-based approach is less salient for more general policies that were not designed to influence a particular research line, but rather shape overall knowledge production or knowledge production. In this setting, the policy would not likely influence specific research lines over others, thus rendering a publication-based approach more suitable. In addition, if a policy shock is likely to move researchers' preferences entirely out of a particular research line and into a different (related or unrelated) area, then the publication-based approach would be more suitable. Moreover, as noted above, researchers must carefully establish that a policy shock is truly unanticipated in order for its impact on research trajectories to be estimated. If the shock were anticipated by the research community or were induced by changes in the nature of scientific progress, then we could not know for certain whether the policy shift caused the subsequent change in knowledge accumulation, whether the post-intervention patterns simultaneously influenced and were influenced by the policy shift, or whether the post-intervention patterns preceded the policy and, thus, any statistical findings attributing post-intervention patterns of knowledge accumulation to the intervention would be specious. Thus, identifying appropriate science policy interventions can be

challenging. Finally, beyond whether a policy is exogenous, and perhaps most salient for our analysis of the hESC policy, it is critical to examine the likely ways in which a particular policy will change researchers' incentives and interests in specific research lines over others. Absent such an understanding, an interpretation of the impact of the policy is extremely challenging.

III. Human Embryonic Stem Cells: Science, Policy & Evaluative Approach

We now turn to the specific setting in which we implement our citation-based approach. We first examine the specific nature of the U.S. stem cell policy intervention and its likely impact on the willingness of researchers to pursue research building on the hESC breakthroughs that arose in the late 1990s. We pay particular attention to whether or not the policy setting is appropriate to our citation-based evaluative approach, particularly highlighting the unexpected nature of the policy change, its influence on follow-on research that builds on the initial hESC discoveries and the opportunities for follow-on researchers to seek resources from sources beyond the U.S. Federal government.

III.1. hESC Scientific Breakthrough

In 1998, James Thomson from the University of Wisconsin and his colleagues published an unexpected scientific breakthrough in leading journal *Science* – the first isolation of human embryonic stem cells.[5,6] Thomson's publication came 17 years after Evans, Kaufman and Martin reported the first isolation of mouse embryonic stem cells (Evans and Kaufman, 1981; Martin, 1981). Although Thomson's lab at Wisconsin had been able to isolate monkey embryonic stem cells two years earlier in 1996, their work with human cells came as a surprise to many scientists and scientific observers. Unlike some other advances, the scientific importance of hESCs was recognized immediately: hESCs were acknowledged both as a critical new research tool for basic research in embryology and as a gateway for applied research that could lead to novel therapies. This dual contribution in basic and applied research positioned the research squarely in Pasteur's Quadrant

[5] Thomson, through the Wisconsin Alumni Research Fund (WARF) also patented his discovery in the U.S. although patents were not granted elsewhere. As a consequence, one important difference between the policy environments in the U.S. and foreign countries is the intellectual property regime. As this feature of the U.S. environment for hESC research does not change contemporaneously with the policy shock, we do not expect that it would confound our analysis or results.

[6] This breakthrough built on a century of research in biology that originated with the observation in the 19th century that certain cells could produce other cells, most notably blood cells. Interest in this line of research accelerated in the late 1960s and 1970s with the development animal based research on in-vitro fertilization techniques, which were propelled by practical considerations regarding fertility, embryology and development. In the 1980s, some physicians and scientists began to extend successes with in-vitro fertilization techniques to humans, while others worked with animal-based embryonic and adult stem cells.

(Stokes, 1997). Indeed, *Science* hailed the work as a "Breakthrough of the Year" in January 1999. The intense interest in stem cells was (and remains) grounded in the fact that "a stem cell is a special kind of cell that has a unique capacity to renew itself and to give rise to specialized cell types. Although most cells of the body, such as heart cells or skin cells, are committed to conducting a specific function, a stem cell is uncommitted and remains uncommitted, until it receives a signal to develop into a specialized cell. *Their proliferative capacity combined with the ability to become specialized makes stem cells unique* [emphasis added]" (NIH, 2001, ES-1). Thus, both the scientific and medical promise of stem cell research derives from its potential to develop multiple types of cells.

The potential of human ESCs as a basis for human therapeutics was also the basis of their controversy – they were derived from human cells and, unlike other forms of "adult" stem cells, they were derived from embryos. The distinction between adult stem cells and embryonic stem cells is central to the hESC debate (and our some aspects of our approach to evaluation): "An adult stem cell is an undifferentiated cell that is found in a differentiated (specialized) tissue in the adult, such as blood. It can yield the specialized cell types of the tissue from which it originated [and…] can renew itself" (NIH, 2001, ES-1). In other words, while it can be transformed into a number of differentiated cell-types, it cannot give rise to *all* types of specialized cells. By contrast, embryonic stem cells are derived from embryos and are pluripotent; unlike adult stem cells, they can develop into any of the over 200 cell types in the body. Therefore, as of the late 1990s and 2000s, leading scientists regarded embryonic stem cells as holding more promise as research tools and as medical therapies because of their greater ability to differentiate leading to considerable interest in and attention on embryonic stem cell research in general. What made the Thomson breakthrough particularly significant and controversial was its heralding of the ability to isolate and manipulate embryonic stem cells from *humans* rather than from other sources, such as mice or monkeys, a step of great significance from a medical perspective and a step that many regarded as likely to form the basis of important follow-on research building on Thomson's results and materials.

Thomson was not entirely alone in focusing attention on hESCs: During the late 1990s, hESC research was taking place in a very small number of international laboratories. Aside from Thomson's laboratory in Wisconsin, leading researchers included a small group in Israel, Australia and the United Kingdom. Perhaps not surprisingly, these groups are not entirely independent of one another. Thomson had a long running collaboration with Israeli scientist Joseph Itzkowitz-Eldor, who, in turn, co-authored a series of hESC papers with Nissim Benvenisty. In Australia, Pera and Trounson worked closely and collaborated with Bongso in Singapore. Indeed, the U.S., Israel,

England, and Canada are the only countries associated with the reprint authors of the "core" hESC research articles listed as seminal in hESC prior to 2001 (see Table 1, Panel C).

Whether building on the research of Thomson or other seminal papers in the field by international scholars, U.S. researchers had to determine whether the policy context would allow them to pursue hESC research and provide funding for their ongoing research.

III.2. U.S. Stem Cell Policy Context: 1970 - 2001

In the decade prior to Thomson's work, there had been considerable debate over the appropriate levels of support for research on human embryos, which were necessary for the isolation of human embryonic stem cells with methods developed by Thomson and others. Beginning in 1973, U.S. government policies prohibited federal funding from supporting research on fetuses, embryos, and tissues associated with either. These restrictions did not, however, impose bans on private sector or privately funded research (Wertz, 2002).[7] What followed was a period with significant policy uncertainty. In 1993, the Democratic-majority Congress passed the National Institutes of Health Revitalization Act, which would have enabled the NIH to distribute funds for research on human embryos. In interpreting the Act, the Clinton Administration offered partial support for research on human embryos, but prohibited the NIH from funding experiments that would create embryos *exclusively* for research purposes. In 1995, however, the Republican-majority Congress passed the more restrictive Dickey-Wicker Amendment, which expressly prohibited federal funding for research in which human embryos were either created or destroyed. This limitation essentially precluded federal support for research on *in vitro* fertilization, which generally creates more embryos than are actually used in fertility treatment, and hence implicitly for hESC research, although this was considered a legal gray area. It was in this context that Thomson himself had received and used corporate funding from biotechnology company Geron to support his research on hESC isolation.

In the two years following Thomson's breakthrough (the final years of the Clinton Administration), the policy environment became significantly more inclined towards federal support for human embryonic stem cell research. And, in August 2000, only a few months before the Bush vs. Gore presidential election, the NIH published guidelines enabling federal funding for research using existing cell lines, and it solicited proposals for future research, which it intended to adjudicate

[7] In the late 1980s, both the National Institutes of Health (NIH) and Congress became more sympathetic to the prospect of using government funding for research on human embryos. Although the NIH's Human Fetal Tissue Transplantation Research Panel voted overwhelmingly in favor of support for funding research involving fetal tissue and human embryos, DHHS Secretary Louis Sullivan rejected the Panel's recommendation and extended the ban on such funding (Wertz, 2002).

in March 2001.[8] Although the prior few years had involve a substantial gathering of momentum for supporting hESC research, the Presidential election led scientists to be uncertain about the policy context they would find in 2001. There were sharp contrasts in the policy proposals on the table from the Democratic presidential candidate Al Gore and the Republican candidate. A closely fought election and contentious outcome, which ultimately resulted in President George Bush as the Republican President continued to fuel the uncertainty. Thus, the United States began 2001 amidst a contentious public debate and an evolving, uncertain policy environment.

The scientific community actively touted the prospective value of hESC research and remained hopeful about the prospects for large-scale funding. President Bush therefore initiated an official review of policy options with respect to human embryonic stem cell research. His administration placed, however, a hold on all funding of hESC proposals solicited by the NIH. In February 2001, as part of the administration's review process, Tommy Thompson, the Secretary of Health and Human Services, requested "that the National Institutes of Health prepare a summary report on the state of the science on stem cells … [which] provides the current information about the biology of stem cells derived from all sources— embryo, fetal tissue, and adult" (NIH, 2001, p. i). The NIH issued its report, "Stem Cells: Scientific Progress and Future Research Directions," later that year, in June 2001 (as we outline in Section IV we will use the research articles described in this document as the source of the core articles describing breakthrough knowledge of hESC in the period prior to the policy intervention).

In August 2001, President Bush introduced his administration's policy. It was met with considerable surprise in the U.S. media and by many scientists.[9] The announced policy involved compromise between opposing viewpoints. Specifically, the announced policy neither involved a full ban on hESC research, nor provided unrestricted funding for hESC research. Instead, the policy offered federal support for hESC research, subject to conditions on research materials. The policy included three features that are notable for our current project: The policy (1) enabled federal funding for research on the set of human embryonic cell lines that were already existed at the time of the policy (including cell lines that were reported in the breakthrough work of Thomson and others

[8] The Clinton Administration's reversal was based on an opinion provided by Harriett Rabb, then General Counsel at the DHHS, to Harold Varmus, then Director of the NIH, in which Rabb concluded that funding research using hESCs that were not derived with federal funds would not violate the Dickey Amendment because hESCs do not meet the definition of embryos (Rabb, 1999; NIH, 1999).

[9] In addition to being a genuine surprise, the Bush policy was met with negative reactions from both the right and left of the political spectrum (Wertz, 2002) and with substantial disappointment within the scientific community (Clark, 2001; McGinley and Regalado, 2002). Among other arguments, opponents of hESC research were dismayed that research associated with the destruction of embryos was permitted under the Bush Administration policy, while proponents of hESC research argued that the limitations on federal funding would inhibit scientific advances and retard medical improvements (Wertz, 2002).

outside the U.S., such as those developed by Trounson in Australia); (2) prohibited federal funding for the development of and research on new human embryonic cell lines; and (3) placed no restrictions on the use of private, state, or local funds for hESC research purposes. Although the policy (a) offered the first unimpeded path to Federal funding of research involving human embryonic stem cells (the Clinton Administration initiated funding requests but did not make grants) and (b) did not place restrictions on private or state-level hESC research funding, the policy constrained scientists with respect to the research materials they could use. Thus, the policy involved targeted research funding, which constrained the autonomy of researchers to choose materials in projects involving Federal grants.

Based on these details, the U.S. hESC policy intervention seems appropriate for evaluation in a citation-based casual framework. First, the policy applies to a specific research area in a way that may alter follow-on research that builds upon a core set of published articles – i.e., the policy directly bears upon follow-on work in hESC research, but does not directly influence work on other areas of research in cell biology. As well, the intervention was expected to be of considerable importance for hESC research, an area expected to have promising implications for medical science. Also of importance for our research design, the policy shock appears to have been exogenous. There is ample evidence of the significant uncertainty in the policy environment surrounding hESC research in the period prior to the policy decision – exacerbated by the uncertainty of the Presidential elections – and the decision itself was also unexpectedly subtle and nuanced. Though the policy intervention appears suitable for analysis in our framework, it could be more ideal. The policy's nuance leaves room for interpretation of whether to consider the shock as one that provides targeted funding for some research areas or imposes funding restrictions relative to the rest of the world. The timing of the shock involves a period of uncertainty (in 2000, which we would expect to have an impact on research output in 2001) and the August 2001 policy intervention itself, which we expect to have an impact on research in the years thereafter.

III.3. Post-2001 Research Context, U.S. and Abroad

From the perspective of researchers with an interest in pursuing hESC research within the United States, a complex variety of options were now available. Researchers interested in pursuing hESC research using approved lines could apply for NIH and other government support to do so. And resources were now available: Of the approximately $550 million devoted to stem cell research in 2005 by the U.S. federal government, about $24 million was devoted to human embryonic stem

cell research (Beardsley, 2005).[10] Researchers interested in applying for NIH grants for support for research using approved hESC lines could do so; however, few of the human embryonic stem cells created before August 2001 are valuable for therapeutic purposes, as the majority were contaminated by mouse embryonic feeder cultures (Martin et al, 2005).

The policy also formally opened an avenue for interested non-federal actors to support hESC research efforts (albeit a complex one from the perspective of funding logistics). Researchers could receive private funding for projects using either approved or non-approved hESC lines. However, those researchers with funding for non-approved hESC lines who also received federal support for research on approved lines were obligated to establish laboratories that were physically and organizationally distinct from one another. Not extensively documented, private funding was available to researchers from both corporate sources as well as from private philanthropic sources. For example, Thomson continued to receive funding from the private biotechnology company Geron for his hESC research at the University of Wisconsin. Moreover, soon after the 2001 policy announcement, leading institutions such as Harvard University initiated major fund raising efforts with private philanthropic donors interested in pursuing the promise of hESC research focus on specific medical applications, such as diabetes. Again, such funding had to be used carefully with strict provisions established around the co-mingling of funds in specific projects and in the laboratory. Precise estimates of annual private hESC research funding are difficult to obtain; based on high-profile public pledges to major universities and the support of for-profit corporations, such as Geron, it appears as if private funding of hESC research approximated that of public research in the years following the policy intervention.[11]

In addition to the private sources of funding described above, the policies enacted by the Federal government left substantial room for both funding and policy choices at the state level. Some states (e.g. Louisiana, South Dakota, Kansas, and Nebraska) enacted restrictive policies. At one extreme, Louisiana established a complete ban on the use or destruction of embryos for any purpose and recognized embryos legally as "juridical persons" (Rewerski, 2007). In general, however, restrictive state bans did not arise in states that were in a strong position to capitalize on Thomson's research and build cumulatively on the hESC agenda. States with a strong record in the life sciences

[10] The NIH reports spending approximately $938 million on overall stem cell research in FY2008, of which $88 million was targeted for hESC research (NIH, 2011). For the sake of relative comparison, the NIH provided $5.5 billion for Cancer research, $3.5 billion for Women's Health, $1.4 billion for Nutritional Health, $400 million for Depression research, $225 for Sleep research, and $80 million each for research on Sickle Cell Disease and Spinal Cord Injury (NIH, 2011).

[11] For example, a series of institutions, including Cornell University, Harvard University, Sloan-Kettering Memorial Cancer Institute, (as well as other universities, such as Johns Hopkins, Stanford, and UCLA) have received pledges from private sources greater than $10 million each (Hughes, 2006).

were among those states that enacted more permissive policies. And, several state governments authorized funding for hESC research or facilities subsequent to the Bush Administration policy announcement (California, Connecticut, Illinois, Maryland, Massachusetts, New Jersey, New York, Wisconsin (Vestal, 2009)). State-sponsored funding did not, however, arrive swiftly. In December 2005, New Jersey became the first state to administer funds for research projects. Even as late as 2007, only a limited fraction of state-approved funding had materialized in the form of research grants (Vestal, 2007).[12]

Yet a third source of resources for U.S. researchers was to seek well-funded collaborators in countries outside the United States whose policy frameworks and government funding enabled joint research on hESC projects using both "mandated" and new human embryonic stem cell lines. As noted above, the international research community in hESC research prior to 2001 was small with only a handful of labs having developed and sustained the expertise to isolate and maintain human embryonic stem cell lines. U.S. researchers therefore had a relatively "thin" albeit expanding research community from which to draw if they sought to find non-U.S. collaborators. Of course, with global collaboration came the need to navigate an extremely complex international policy context, to understand the global community of hESC researchers, and to find a division of labor that allowed hESC work to take place outside the U.S. with complementary efforts inside the U.S. labs.

In the post-2001 policy period, a number of European countries - including Austria, Ireland, and Italy - did not permit scientists to derive stem cell lines, conduct research on existing lines, or conduct research involving somatic cell nuclear transfer. Germany allowed research on existing lines, but prohibited nuclear transfer and the derivation of new lines. Countries with more permissive policies, however, included a number of leading research nations: Israel, Singapore, Sweden, and the United Kingdom, as well as some of the rapid followers such as China, Japan, and South Korea (Walters, 2004). Indeed, nations including the U.K., Israel, and several Asian nations generally increased their commitment to hESC research and hoped to gain a comparative advantage in hESC research. However, the level of funding was highly limited compared to the (albeit limited) levels available in the U.S. Although it would be attractive to explore the impact of specific national policies on country-specific hESC research output, it does not appear as if the data and phenomenon support our ability to do this, as we do not have well-defined policy interventions outside the United

[12] The U.S. Administration Stem Cell policy was altered in March 2009 when President Barack Obama issued an executive order overturning the ban on the use of federal funding for deriving new human embryonic stem cell lines and conducting research lines derived after August 2001. The NIH authorized approximately $20 million in funding for human embryonic stem cell research in 2009 and approved the first set of new hESC lines in December 2009. Our data do not, however, cover this new period of policy, however, as these changes would be too recent to analyze with confidence. Since the Obama Administration's order in 2009, a series of additional political and legal actions have created further uncertainty regarding the extent of future hESC funding in the United States.

States (a) that involve unanticipated shocks and (b) that affect a sufficiently large segment of researchers that would enable us to identify the causal impact of national policies on hESC production. These limitations also inhibit our ability to identify the impact of the U.S. policy shock on hESC research output in individual countries other than the United States.

Our descriptive statistics will illustrate (see Section V) among citations to the core hESC research articles prior to 2002 (i.e. citations to hESC articles in the pre-policy period), the U.K., Germany, Australia, Japan and Israel dominate. In later periods, researchers from China, South Korea and Singapore also start to make significant contributes to the area of hESC research, in line with both growing investments in science overall but also with their open policies towards hESC research. It is, however, important to note two additional facts about hESC research production and policy outside the United States that are particularly critical for our research design: (1) the production of hESC in countries other than the U.S. was not highly concentrated (Levine, 2005) and (2) while the policy landscape was dynamic, few countries undertook radical changes in their hESC policies contemporaneously with or as a result of the Bush policy and there were no large-scale policy shifts synchronized across a number of countries that could be thought to interfere with our policy analysis.

III.4. Predictions regarding the impact of the hESC policy intervention

We interpret the Bush Administration policy shock as involving two related elements. First, we consider the period of uncertainty around the 2000 Presidential election and the period preceding the announcement of the Bush Administration policy as constituting a shock relative to the expectations regarding future hESC research funding building during the Clinton Administration. We expect that this shock would have a negative impact on U.S. hESC research *relative* to other countries that did not experience this uncertainty during this particular year. Second, we consider the implementation of the Bush Administration policy beginning in August 2001 as constituting an intervention that shifted the attractiveness of hESC research in the United States (in comparison to other fields) relative to the rest of the world (i.e., making the prospect of hESC in the U.S. less attractive relative to the pre-Bush period, and hence, relatively less attractive to U.S. researchers than non-U.S. researchers who did not experience a shock in this particular period). Although the policy ensure that Federal funding would, indeed, by available for hESC research, we expect that the targeting of the funding to a restricted set of research materials, would yield a reduced the level of participation in human embryonic stem cell research in the United States relative to the rest of the world, in which such research fund targeting was not enacted beginning in that period.

Specifically, in the post-August 2001 period, U.S.-based scientists were faced with *relatively* higher difficulty accessing particular hESC materials; thus, we anticipate that scientists faced with these constraints would be relatively less likely to participate in hESC research, a prediction in line with recent studies of the role of biological resource centers in increasing participation in research lines associated with easily accessible materials (Furman and Stern, 2011).[13] This prediction is also consistent with recent literature emphasizing the importance of research freedom to scientists, and with earlier research that emphasizes the importance of field size and legitimacy for continued researcher commitment (Garud and Rappa, 1995).

While our first prediction is simple in its formulation, recent studies suggest additional predictions that would highlight the factors that shape scientists' responses to the hESC intervention. Specifically, qualitative studies have countered the view that scientists are passive recipients of the organizational rules, incentives and governance systems imposed by those who fund them, and reports have challenged the idea that scientists simply conform to any changes imposed on them. Indeed, there is evidence to suggest that the very autonomy and strong normative nature of the scientific community allows scientists to engage higher levels of strategic action than previously anticipated. For example, Murray (2010) recounts the response of the "mouse community," a community of scholars that employ mice in research, to Dupont's exercise of strong patent rights over mice bred with a genetic predisposition to cancer. Qualitative evidence documents mouse researchers' attempts to subvert or ignore the intrusion to their research freedom and suggests that those scientists with greater resources take higher levels of strategic action. In other words, when confronted by expanding restrictions on their use of research inputs, scientists construe these actions as intrusions on their scientific freedom and take strategic action to respond and resist, with the aim of pursuing the research paths that they perceive to be most valuable.

Accordingly, we predict that the higher status scientists, who possess greater access to both financial capital and social capital, are less likely to be affected by policy targeting. Instead, over time, they would use their resources to buffer themselves from policies that attempt to steer their research choices in a direction not of their own choosing. During the period of our analysis, new hESC research lines were perceived by scientists to have greater scientific potential, on average, than the lines approved for funding under the Bush Administration. Thus, we expect that researchers with the status or other means to do so would identify ways to continue participating in hESC research efforts. Specifically, we predict that scientists at high status universities and those with international

[13] The popular press worried openly about the possible negative impact of the policy on U.S. competitiveness in hESC research (Clark, 2001; Anand and Regalado, 2002) and similar concerns were expressed by members of the scientific community in leading journals such as *Science* (Vogel, 2001) and the *Nature* journals *(*Smaglik, 2001; Fletcher, 2001).

collaborators will find strategies to circumvent the controls on research direction, and thus exhibit higher levels of research participation, lowering the impact of funding boundaries on their levels of participation in hESC research.

Taken together, our predictions suggest that the 2001 hESC policy will have a significant impact on U.S. scientists attempting to build upon the important research trajectories established in hESC research in the period from 1998 to 2000. This impact will be one of a significant decline, relative to their foreign counterparts and relative to expected trends in the globalization of research along high impact research trajectories. In addition, we predict that there will be a differential impact on different groups of U.S scientists, again relative to expected trends. To evaluate these questions, we use the citation-based evaluation approach outlined above. In the section that follows we describe our methods paying particular attention to the development of the core set of (treated) hESC research articles and the different groups of control articles. This is essential because our results are predicated not only on capturing differences in research trajectories pre- and post- the policy intervention but also on evaluating those differences relative to a well-defined counterfactual.

IV. hESC policy evaluation: Design, estimation and data

IV.1. Research design

Our methodological approach to evaluating the 2001 hESC policy follows the critical steps laid out more generally in Section II. We take the core set of hESC research articles from 1998-2000 and use the core set of RNAi research articles during the same period as our primary counterfactual. We then implement a difference-in-differences analysis to estimate the impact of the 2001 hESC policy intervention on the trajectory of human embryonic stem cell research (i.e. comparing the pre- and post- policy citation trends) relative to that of our RNAi controls for U.S. researchers. The citation-based approach we follow here differs from the publication-based and keyword oriented approach used in prior policy evaluations of the hESC intervention. In particular, Owen-Smith and co-authors (Owen-Smith and McCormick, 2006; McCormick, et al., 2009; Scott, et al., 2009, 2010, 2011)) as well as Levine (2005; 2008) and Löser et al. (2009) have explored the flows of scientific knowledge in the stem cell arena tracing out the patterns of stem cell research (using keywords) over time and across geographies, thus providing important insights into the general patterns of hESC and other stem cell research overall. The methods used in these studies provide key insights into broad patterns of hESC research. In contrast, our citation-based approach focuses very explicitly on what happens to follow-on research that builds on the core hESC breakthroughs after the policy shift that changed the conditions of and incentives for these trajectories. There are clearly benefits as well as costs to the citation-based methodology in the context of the hESC policy intervention. Unlike the

keyword approach, we cannot capture research that establishes *entirely* new and novel approaches to the production and isolation of hESCs, i.e., to totally new and unconnected research lines. In other words, this approach is very poor at evaluating policy shifts that move research to entirely new fields of enquiry or to exploring individual researchers' choices of disconnected arenas. In contrast, however, we can provide a much more precise estimate of the causal impact of the policy on the hESC trajectory overall and more specifically on U.S. researchers in different types of institutions.

A core element of our research design is the identification of a plausibly exogenous shock to cumulative research in a particular trajectory. In this case, the shock involves the resolution of the 2000 U.S. presidential election and the subsequent 2001 announcement of the Bush Administration's hESC funding policy. Considering this as the policy shock around which we build our analysis involves the assumption that the policy intervention constituted an exogenous shock suitable for analysis in our framework. As we describe in Section III.4, we believe that the historical record suggests this to be a reasonable assumption.[14]

Accordingly, we conceive of the U.S. hESC research landscape as consisting of the pre-election period, the period between the announcement of 2000 election results (in December 2000), and the period following the announcement of the Bush Administration policy (in August 2001), during which the policy intervention may influence the rate and nature of cumulative hESC research efforts. Considering that the delay between research project initiation and publication is approximately six months to a year, we believe that it makes sense to consider the pre-2001 period as the "pre-intervention" period in the analysis, the year 2001 as the year in which the results of the election may have an impact on hESC publication, and the years after 2001 as the period in which the impact of the announced policy can be ascertained.

Having defined a policy intervention that may affect cumulative knowledge, the next step in applying our methodological approach involves identifying a clear control group that can help provide a counterfactual estimate of the trajectory of follow-on hESC research. The trajectory of research in RNA interference (RNAi) constitutes our primary control group, although we examine two additional potential controls in our robustness checks.

[14] The hESC research community had reason to expect that the Bush Administration may not be as supportive of hESC research as was the Clinton Administration. This was borne out by the NIH's ceasing evaluation of previously-submitted hESC proposals, which effectively reversed the course followed towards the end of the Clinton Administration, which had announced expectations regarding funding hESC work. The resolution of the uncertainty seemed favorable relative to expectation that hESC research may be forbidden by the Bush Administration. Indeed, the Bush policy mandated the first federal funding for hESC research. However, the condition that only pre-existing cell lines would be eligible for federal support was interpreted by many in the scientific community as a restriction on funding for research materials relative to their expectations prior to the Bush electoral victory.

RNAi research has several important characteristics that allow it to serve as an appropriate control group. It represents a scientific breakthrough that was (a) pioneered in the United States achieved, (b) in essentially the same field and at the same time as hESC research (1998), and (c) perceived to be of similar scientific importance in 2001. Thus, hESC research and RNAi research share similarities in that both were life science breakthroughs, pioneered in the U.S. in 1998, and that each diffused to the rest of the world at similar rates prior to the 2001 policy intervention. It is also helpful that, unlike some other areas of research, including stem cell research involving adult stem cells or animal models, methods in RNAi and hESC are relatively independent. It is difficult for researchers to simultaneously conduct or to substitute between RNAi and hESC research and, according to our qualitative evidence, researchers do not use both methods in their laboratories. Hence, a policy shock affecting the extent and organization of U.S.-based hESC research is not likely to have a direct impact on RNAi research. Moreover, unlike hESC research, RNAi was not subject to any specific policy intervention during the time of our analysis. Based on similar considerations, Levine (2008) also argues for the suitability of RNAi as a comparison field for hESC.

The next step in implementing our approach is to link the treated and control groups with a set of reference articles that reflect core knowledge in each area. To do this for hESC research, we rely upon a sample of stem cell articles identified by the NIH report, "Stem Cells: Scientific Progress and Future Research Directions." This report was published in June 2001 and was developed to be an input into the Administration's policy-making process. The report was devoted to scientific facts relevant to the policy debate, but does not appear to be a political document.[15] Importantly for our research design, the document identifies 110 articles deemed by the NIH to reflect the seminal articles in stem cell research, including papers associated with embryonic and adult stem cells derived from both human and animal models. We consider the 17 hESC articles to be our core articles – our primary treatment sample. (Of these, one article is not referenced in the ISI Web of Science data set, so our primary hESC sample involves citations to the 16 remaining articles.) To identify the core research on which the RNAi research trajectory builds, we rely upon the list of seminal RNAi articles published by Ambion Inc., a company that manufactures and markets products related to RNAi-research. The list includes 56 articles, of which 52 were published prior to or during 2001 and 4 of which were published in 2002.[16]

[15] The report notes in its Preface, "NIH recognizes the compelling ethical and legal issues surrounding human pluripotent stem cell research. Because extensive discussions regarding these issues have been presented in various forums elsewhere, they are not part of this review of the state of the science. Also, the report does not make recommendations pertaining to the policies governing federal funding of such research" (NIH, 2001, p. II).

[16] Our results are not sensitive to including or omitting the four RNAI articles published in 2002.

While considering RNAi core articles as our primary control sample, we do examine two alternative control samples, which we include in data comparisons and in robustness checks, although we do not consider them to be ideal controls. These samples include (i) a set of matched articles appearing in the same journals at the same time as the hESC treatment articles, which we call "nearest neighbor" articles (following Furman and Stern, 2011) and (ii) areas of stem cell research that involve either adult stem cells or animal stem cell models (i.e., stem cell research other than hESC). The nearest neighbor sample consists of the three articles that immediately proceed and follow each of the root hESC articles in the same year and issue of the journal in which the hESC root article was published. The nearest neighbor articles are, therefore, precisely matched with the hESC articles with respect to publication timing and journal. They may, however, be poorly matched with respect to scientific field; this is particularly apt to be the case for articles that appear in general interest science journals, such as *Science* and *Nature*, in which articles in life sciences may be published contiguously with articles in physics or even social sciences. The nearest neighbor articles are also less likely to be matched in scientific importance. Thus, we do not consider the nearest neighbors to be an ideal control sample, as we feel believe that it is more reasonable to rely on comparisons to another "hot" area of science like RNAi than to areas of "normal science," as these types of research may be subject to different dynamics (Kuhn, 1962).

Stem cell research using adult or animal models ("Other Stem Cell" research) is intuitively appealing as a control group as it bears direct similarity to hESC research and because the Bush Administration hESC policy did not alter funding for these areas of research. In addition, 93 seminal OSC reference articles are identified by the same NIH Report from which we draw our hESC reference articles. After considerable discussion with scientific expertise, however, we decided against using OSCs as our main control group because the extent of research in this area was likely affected by the Bush Administration policy in an important, though indirect, way. When researchers interested in pursuing hESC research faced a limited set of funding options and examined alternative research paths to pursue, the use of animal ESCs or adult stem cells were natural alternatives. Indeed, many laboratories at the time (and today) used a combination of stem cell methods and materials. Consequently, by the 2001 policy intervention would likely involve *substitution* between hESC and OSC research. As a result, OSC research does not constitute an independent control that could establish a clear counterfactual to hESC research, but is more similar to another "quasi-treatment" group, subject to a different but related treatment.

IV.2. Estimation

The final element of our research approach involves econometrically estimating the causal impact of the 2001 policy on U.S. hESC research relative to Non-U.S. hESC research. The intuition behind the approach is that draws two simultaneous comparisons. We (a) compare citations to core hESC articles from U.S.-based reprint authors vs. citations from Non-U.S. reprint authors before and after the policy intervention and (b) compare these trends to differences in citations to core RNAi articles from U.S.-based and Non-U.S.-based reprint authors. The comparison of U.S. to Non-U.S. RNAi citations establishes a baseline relative to which we can then compare trends in U.S. and Non-U.S. hESC citations. In principle, this is similar to the descriptive analysis by Levine (2008), which compares country-level counts of citations to Thomson et al.'s (1998) seminal hESC paper with those to Fire et al.'s seminal RNAi paper, both before and after the policy intervention. The central advantages associated with using econometric analysis, rather than straightforward counts, are that the differences-in-differences techniques we employ are able to identify fine-grained trends, including year-to-year trends, in relative hESC and RNAI research output and are able to provide precise estimates (i.e., percentage increases or decreases) of the causal impact of the hESC policy intervention on post-intervention output in the United States.

Estimating the impact of the policy intervention econometrically is complicated by the fact that the analysis requires two different dependent variables, one reflecting citations by U.S.-based and another citations by Non-U.S. based reprint authors. One way to do this would be to run these models separately, including the same dependent variables in each model and then comparing the coefficients indicating the post-policy impact on each. A drawback of this approach is that it will, necessarily, assume that the set of factors driving U.S. and Non-U.S. citations is different. This makes it difficult to compare coefficients indicating the impact of the policy intervention on citations.

To overcome this difficulty, we simultaneously estimate the drivers of U.S. and Non-U.S. citations in a single equation, using a single dependent variable, $CITES_{it}^{r}$, with a data structure that involves replicating the entire dataset twice by "stacking" the data upon itself. The top half (top stack) of the resulting dataset differs from the bottom half (bottom stack) in one key way. In the top half (stack) of the data, the variable $CITES_{it}^{r}$ reflects citations generated by U.S.-based reprint authors (i.e., $CITES_{it}^{r=US}$), while, in the bottom half (stack), it reflects citations by papers whose reprint authors are from countries other than the United States (i.e., $CITES_{it}^{r=NotUS}$). The superscript *r* indexes whether the data are from the "U.S. stack" or the "Non-U.S. stack" of the data. The resulting dataset also includes a stack-specific dummy variable indicating that the top half is the U.S. stack and that the bottom stack is the Non-U.S. stack. With these exceptions, however, the top and bottom stacks include exactly the same data. Organizing the data in this structure enables us to

identify the impact of U.S. policy on citations generated by U.S.-based authors relative to the impact of these policy periods on citations generated by authors based in other countries, holding all other variables constant that simultaneously affect citations in the U.S. and the rest of the world. This approach follows that of Furman and Stern (2011) and Murray et al. (2011).

With this data structure in place, we estimate:

$$CITES_{it}^{r} = f(\varepsilon_{it}; \gamma_i + \beta_t^r + \delta_{t-pubyear}^r + \alpha_0 HESC_i * 2001_{it} + \alpha_1 HESC_i * (t > 2001_{it}) + \psi_0 (US_i^r * HESC_i * 2001_{it}) + \psi_1 (US_i^r * HESC_i * (t > 2001_{it})) \quad (1)$$

where the superscript *r* denotes the region to which the variable pertains (r = either U.S. or Non-U.S., depending on the address of the article's reprint author), *i* indexes each article and *t* indexes each year, while (γ_i) is a fixed effect for each article, β_t is a year effect, $\delta_{t\text{-pubyear}}$ captures the age of the article, and *HESC*2001* and *HESC*(2002-2007)* represent dummy variables equal to one for hESC articles in years 2001 and 2002-2007, respectively.[17] The coefficients on these variables (α_1 and α_2) identify the difference in follow-on research experienced by hESC core articles (both associated with U.S. and Non-U.S. reprint authors) relative to the control articles during the 2001 and 2002-2007 time periods. The coefficients ψ_0 and ψ_1 are the central focus of our analysis. These indicate the marginal impact of the policy intervention on follow-on research in the United States relative to the rest of the world and relative to follow-on research in the control group. In other words, these coefficients indicate the additional increment or decrement to citations that hESC root articles receive in the U.S. relative to the rest of the world in the year of and the years subsequent to the policy intervention.

The specification includes one key parametric restriction: We estimate the reference article fixed effects (γ_i) to be equal in each stack of the data. We allow calendar year and article age fixed effects to vary by stack, however. By estimating the stacked regression, rather than estimating the drivers of CITES separately for the U.S. and Non-U.S. authors, we are able to interpret ψ_0 and ψ_1 as indicators of the change in hESC article output in 2001 and the years after the policy shift relative to hESC article output in the rest of the world.

[17] Several issues, including the incidental parameters problem, arise in incorporating multiple fixed effect vectors into a negative binomial specification. We have experimented with a range of alternative procedures and approaches, including the conditional negative binomial estimator suggested by Hausman, Hall, and Griliches (1984) and the fixed effects estimator suggested by Allison and Waterman (2002). Our core results are based on the traditional conditional fixed effects estimator with bootstrapped standard errors; however, our qualitative findings are consistent across these different procedures. When using a conditional fixed effects estimator, one citation year and one age fixed effect are not separately identified (Hall, et al., 2007).

We further modify (1) to accommodate more flexible ways of identifying the impact of the policy intervention. In particular, we estimate:

(2)

$$CITES_{it}^{r} = f(\varepsilon_{it}; \gamma_i + \beta_t^r + \delta_{t-pubyear}^r + \alpha_0 HESC_i * 2001_{it} + \alpha_1 HESC * (2002-2003_{it}) + \alpha_2 HESC * (2004-2007_{it}) + \psi_0 (US_i^r * HESC_i * 2001_{it}) + \psi_1 (US_i^r * HESC_i * 2002-2003_{it}) + \psi_2 (US_i^r * HESC_i * 2004-2007_{it})$$

and

$$(3)\quad CITES_{it}^{r} = f(\varepsilon_{it}; \gamma_i + \beta_t^r + \delta_{t-pubyear}^r + \sum_{t=1996}^{2007} (\alpha_t HESC_i * YEAR_{it}) + \sum_{t=1996}^{2007} (\psi_t US_i * HESC_i * YEAR_{it})$$

where (2) decomposes the post-Bush policy period into effects associated with 2002-2003 and 2004-2007 and (3) decomposes the pre- and post-policy periods into individual year*policy environment fixed effects. The results of (2) will allow us to understand whether the post-policy effects are constant over time. Estimating (3) enables us to estimate the year-by-year effects of the policy intervention and to check for the presence of a pre-policy time trend. The former is important to understanding the dynamic consequences induced by the policy intervention – for example, whether the impact of the policy intervention occurs as a one-time change in the level of or diffusion of knowledge, whether it declines in relative terms or returns to baseline over time, or whether the policy intervention induces continuously growing effects. The latter is important, as an understanding of the pre-policy time trend can provide evidence about the exogeneity of the policy intervention itself.

In addition to examining the impact of the policy intervention on the extent of follow-on research in the U.S. and abroad, we are also interested in whether the policy shift affected the organization of hESC research. In particular, we investigate whether the policy changes had a differential impact on elite and non-elite U.S. universities and on the geographic nature of research collaborations. To estimate the impact of policy interventions on these subpopulations, we decompose annual citation counts by each subpopulation. Specifically, in these analyses, we estimate a version of (2) in which we have three rather than two stacks of data. When investigating the impact of the policy shift on elite and non-elite universities, the dependent variable in the first, second, and third stacks of the data reflect citations to root articles by (a) Top 25 U.S. universities, (b) Non Top 25 U.S. institutions, and (c) institutions outside the United States, respectively.[18] We

[18] We base our definition of elite universities on those classified as being in the "Top 25" by the Center for Measuring University Performance at Arizona State University's 2006 Annual Report of University Research Rankings. (Our

use this approach to investigate the impact of the policy shock on U.S.-international collaboration, by estimating a version of (2) in which the first, second, and third stacks of the data reflect citations to root articles by (a) authors from U.S. institutions only, (b) authors from both U.S. institutions Non-U.S. institutions, and (c) authors from Non-U.S. institutions only, respectively.

V. Empirical analysis

V.1. Descriptive Statistics

Our complete dataset includes 719 root articles, of which 16 are human embryonic stem cell root articles, 51 are RNAi root articles, 559 are nearest neighbor articles and the remainder are 93 are other stem cell root articles. We report descriptive statistics for the root articles in Table 1. Panel A presents data for the entire sample of root articles; Panel B decomposes the data by article type (hESC, RNAi, nearest neighbor, and Other Stem Cells); and Panel C identifies the number of reprint authors by the country of origin for each of the reference groups. Across the sample, more than 60 percent of root articles include a reprint author based in the United States. More than two-thirds of Reprint Authors' addresses identify university affiliations; 24 percent of Reprint Authors are associated with "Top 25" universities, according to a classification scheme based on the Center for Measuring University Performance (Arizona State University) 2006 Annual Report of university research rankings. Nearly half of the root articles involve papers with only U.S.-based authors; 13 percent involve collaborations between U.S. and Non-U.S. authors; and 39 percent of root articles involve only authors based outside the United States.[19] Panel C demonstrates that root articles are concentrated in the United States in each of the sample sub-groups. The hESC root articles include reprint authors from Israel (4), Canada (2), and England (2) as well as the United States (5). By contrast, the nearly three-quarters of the RNAi root articles include U.S.-based reprint authors. If the number of pre-policy citations to these root articles were to reflect a similar fraction of U.S.-based authors, it would imply that the pre-policy intervention RNAi and hESC samples were poorly matched.

Table 2 reports bibliometric characteristics for annual citations received by our root article sample. Panel A presents the data for the entire sample, while Panel B decomposes the data by

---

results are also robust to using the "Top 50" as defined by the same report.) Articles are classified into these institutional categories based on the addresses of their reprint author.

[19] One issue worth noting in Panel B is that the fraction of U.S.-based authors is higher among RNAi root articles than hESC root articles. If RNAi root articles received more citations from U.S.-based authors than did hESC root articles in the period after the initial discoveries enabling RNAi and hESC research (1998-2000), we may be worried that the initial geographic distribution of discovery drove the subsequent geographic distribution of follow-on research. As the data we present below shows, however, the geographic distribution of RNAi and hESC research follows similar patterns between 1998 and 2000.

article type. On average, articles in our sample receive 17.3 citations per year, of which slightly less than half involve a US-based reprint author and only a small fraction of which involve collaboration between U.S. and Non-U.S. co-authors. The standard deviation in Annual Citations is 39.4, which highlights the extent to which these data are skewed. Similar to the root articles, the overwhelming majority of citations involve authors whose addresses can be linked to universities. A smaller fraction of the citing articles, however, are associated with reprint authors affiliated with Top 25 U.S.-universities. There are some noteworthy differences across article types. Whereas hESC root articles receive, on average, approximately 32 citations per year, RNAi root articles receive nearly 68 citations. In addition, RNAi articles receive more citations from U.S.-based authors. The nearest neighbor sample is the least well-cited, receiving, on average, fewer than 9 citations annually. This highlights the fact that the nearest neighbor articles are more representative of "normal science" than the RNAi or stem cells samples, each of which reflects a "hot" scientific field. Panels C and D of Table 2 indicate the country of origin of citing article reprint authors for the periods before 2001 and after 2001, respectively. The fraction of citations from U.S.-based authors in the pre-intervention period is similar in RNAi sample, which helps alleviate the worry that the initial geographic distribution of discovery drives the subsequent geographic distribution of follow-on research. By contrast, the ratio of U.S. to Non-U.S. citing articles in the RNAi sample is higher in the post-intervention period than it is in the hESC sample. This fact about the raw data on citations is consistent with that found by Levine (2005, 2008). In addition to the specific U.S. vs. Rest-of-the-World comparisons, Panels C and D indicate the general research strength of Japan, Germany, and England in each area of science in our sample, as well as the relative strengths of Israel, China, and South Korea in hESC research and China in RNAi research. Additional evidence regarding the strength of these countries in hESC and RNAi is found in Appendix Table 1, which documents the number of papers on which U.S. authors collaborate with authors from other countries. The strength of specific research institutions is presented in Appendix Table 2, which lists the count of follow-on hESC by the Reprint Author's institutions. This table highlights the strength of Johns Hopkins, Wisconsin, Harvard, Stanford, and Geron in the United States and institutions outside the U.S., including Israel research organizations, such as Hebrew University, the Technion, and Rambam Medical Center, as well as Australia's Monash University, the University of Sheffield, Kyoto University, and the National University of Singapore.

V.2. Publication Trends

Figure 1 depicts the number of citing articles by broad article type and year, not distinguishing by country-of-origin. There is an upward trend among each of the subsamples. With

the exception of the hESC sample, the rate of follow-on publications accelerates noticeably between 1998 and 2002 (or 2003, depending on the sub-sample). The raw number of nearest neighbor citations is greatest. This is not surprising, as the baseline number of nearest neighbor root articles is 400 more than that of any other sub-sample. Prior to 1999, the number of articles building on hESC and RNAi roots is relatively similar. Beginning in 2000, however, the extent of cumulative research in these two areas diverges appreciably, as citations to the RNAi roots rise from fewer than 1,000 in 2000 to more than 5,000 by 2004.

Figure 2 reports citing articles by publication type and year for U.S. and non-U.S. reprint authors. It includes four separate graphics, one for each of our samples. Each graphic reports the number of citations to a different root article sample by papers with either (a) any U.S.-based author or (b) no U.S.-based author. In each case, the number of overall citations rises before falling in the final years of the sample. The fact that the overall number of citations declines in the final few years of each graphic is reflective of a typical citation pattern, in which root articles receive the highest number of citations in the few years after their publication (Furman and Stern, 2011). The rate of obsolescence is of less interest in our analysis than the relative levels of U.S.-based and non-U.S.-based citations.

The top-most graphic in Figure 2 compares trends in citations to hESC root articles by U.S.-based and non-U.S.-based reprint authors. While the number of citations by each category is similar between 1998 and 2000, the counts diverge in 2001. Specifically, while the growth rate of non-U.S.-based citations continues after 2001 (until 2004), the relative number of US-based citations declines beginning in 2001, although there appears to be a modest recovery in 2004. These findings are consistent with those reported by Owen-Smith and McCormick (2006), who conclude based on a keyword approach to identifying hESC publications up until 2004, that the U.S. share of hESC articles experienced a relative decline beginning in 2001. The U.S. share of Other Stem Cell citing articles (i.e., articles citing "root" articles in areas of stem cell research other than hESC research) declines beginning in 2003, though not before. U.S.-based citations to RNAi root articles and nearest neighbor root articles also experience a relative decline in the later years of the data, but do not experience a relative decline between 2001 and 2003.

Interpreting these trends requires care and more structured analysis. In light of the relatively stable U.S. share of RNAi and nearest neighbor articles, the unambiguous relative decline of the U.S. share of hESC articles and modest relative decline in other stem cell articles may suggest that the shock to U.S. funding policy had an impact on the rate of follow-on stem cell research in the United States. Consistent with the expectation that the 2000 election created substantial uncertainty about the future of hESC research in the United States, the relative decline in U.S. hESC share begins in

2001. Consistent with the expectation that the U.S. policy did not enhance national hESC output, the relative decline appears to continue in the years after 2001 as well. In order to identify the impact of the election and policy shock precisely, however, we employ the difference-in-differences techniques we describe above.

V.3. Core Results

Our descriptive statistics suggest the importance of measuring hESC research output relative to a carefully matched comparison group. In addition, they highlight the importance of controlling for the impact of timing and root-article effects on follow-on research. In our regression analyses we attend to these issues mainly by relying upon RNAi research as our control group. Thus, we derived our counterfactual estimate of the impact of U.S. hESC policy by comparing geographic trends in hESC output to those experienced by RNAi research. This seems to be an apt comparison, as both areas of research were pioneered in the United States in 1998; both by individuals at institutions of similar relative status, and both are in the same broad area of cell biology. In our robustness checks, we compared the geographic evolution in hESC research to that of research using Other Stem Cells and science overall, based on a sample of nearest neighbor articles. A principal concern with using Other Stem Cells as a control group is the possibility that hESC researchers could substitute into OSC research (and vice versa). Techniques associated with research on human embryos could be applied to mouse embryos, for example. The prospect of substitution between hESC research and RNAi research is unlikely (and citations are not overlapping), however, as the techniques and materials vary greatly. The nearest neighbor articles also form a less satisfactory control group, as publications within a journal often vary in nuanced ways, even across subfields, thus, may have substantially different citation profiles.

We present our regression results beginning in Table 3 with a conditional fixed effects negative binomial specification in which we estimate $CITES_{it}^{r}$ as the dependent variable with two "stacked" equations. In the top stack, $CITES_{it}^{r=US}$ reflects citations by papers with reprint authors in the United States in the top stack of the data; in the bottom stack $CITES_{it}^{r=Not-US}$ reflects citations by papers whose reprint authors are from countries other than the United States. We do not report the significance of tests of joint restrictions on the article fixed effects, as these are not computed in conditional fixed effects models. In addition to reporting raw coefficients and associated bootstrapped standard errors (MacKinnon, 2002), our tables report the coefficients in our results as incidence-rate ratios (IRRs), which are easily interpreted as percentage changes relative to a baseline

(i.e. the null hypothesis of no effect yields a coefficient of 1.00, while a coefficient equal to 1.50 implies a 50% boost to FORWARD CITATIONS).

Column (3-1) estimates the impact of U.S. policy on citations in year 2001 and in the average of the years following the shock (2001-2007). Table 3 reports coefficients that reflect the relative output of hESC vs. RNAi research and U.S. vs. Non-U.S. hESC research. Specifically, the coefficients on HESC*2001 and HESC*(2002-2007) indicate the boost in worldwide hESC output relative to RNAi output during the time periods 2001 and 2002-2007. These are not statistically different from one. The coefficients on US*HESC*2001 and US*HESC*(2002-2007) indicate the boost (or decrement) in U.S.-based hESC output during these periods relative to research output in countries other than the United States. The magnitudes and significance levels of these coefficients imply that, relative to hESC research outside the United States, the production of follow-on research with any U.S. address declined severely beginning in 2001. Indeed, in this specification, the impact of the shock is greatest in 2001, during which research output falls by slightly more than 51%. In the years after the shock, the amount of hESC follow-on papers with a U.S.-based reprint author declines, on average, by 41%.

By estimating the average treatment effect across the years 2002-2007, the specification in (3-1) masks some of the interesting dynamics associated with U.S.-based hESC research. In (3-2), we decompose the post-policy time period into effects associated with the years 2002-2003 and 2004-2007. In the immediate years after the implementation of the Bush Administration policy, U.S.-based hESC follow-on work declines by nearly 59% relative to non-U.S. based research. However, during the period 2004-2007, the production of hESC follow-on papers is only 29% lower in the U.S. than the rest of the world. We consider this dual finding – i.e. of initial relative decline in U.S. hESC research output followed by a research rebound – to be one of the core findings of our analysis. Specifically, (3-2) suggests that the impact of the Bush Administration hESC was greatest in the first few years after its announcement and implementation, but that U.S. hESC output began to recover beginning in 2004, only a few years after the policy was enacted. Although some state funding did begin to support hESC research in the U.S. during the 2004-2007 period, little of it had arrived by the time that the rebound had begun.

We explore the year-to-year dynamics of the relative decline and partial rebound in U.S. hESC research output in the analysis in (3-3) and turn thereafter to an examination of some of the mechanisms of the recovery. Specifically, in (3-3) we estimate a more flexible specification, in which we report year-by-year effects both for all hESC articles (relative to RNAi articles) and U.S.-based hESC articles (relative to articles produced outside the U.S.). Figure 3 plots the year-by-year hESC*U.S. incident rate ratio, along with the associated 95 percent confidence intervals. The results

of this exercise underscore the findings reported in (3-2): U.S.-based hESC output experiences a penalty relative to Non-U.S. hESC output in 2001, 2002, and 2003. At its relative nadir in 2003, U.S. hESC output was only 40 percent of Non-U.S. output. U.S. hESC output recovers between 2004 and 2007, reaching levels that are approximately 66 to 74 percent of Non-U.S. output in 2004 and 2006, and to levels that are statistically indistinguishable from Non-U.S. output by 2007. These results paint a picture that suggests that U.S. hESC research output experienced a substantial decline relative to its potential in the year in which the Bush Administration policy was enacted and the two that followed. However, this decline is followed by what appears to be a steady rebound, which continues until the final year of our data, 2007.

In Table 4, we show whether these results are an artifact of our particular sample and means of classifying national origin. Column (4-1) replicates the analysis of (3-2), using an alternative measure of location in which citing articles are considered to be U.S.-based if any address in the list of author affiliations is based in the United States. These results are qualitatively and quantitatively similar to those of (3-2). Specifically, we find both an immediate relative decline in U.S.-based output in 2001, which continues in 2002 and 2003, along with a recovery between 2004 and 2007. Equation (4-2) reports the result of the same specification as (3-2), again using reprint author to identify the location of the paper; however, rather than comparing hESC output to the RNAi controls, the nearest neighbor articles are substituted for the controls. The positive and significant coefficients on the HESC*YEAR variables suggest that hESC root articles received 59 percent, 92 percent, and 160 percent more citations in 2001, 2002-2003, and 2004-2007 than did their nearest neighbor root articles, thus confirming that hESC research was a “hot” research area relative to normal science. The coefficients on US*HESC*YEAR compare the output of follow-on hESC articles in the U.S. to those outside the U.S., relative to geographic trends in citations to the nearest neighbor articles. These results are similar to those found using the RNAi control sample. Once again, U.S. hESC output experiences a relative decline in 2001 and 2002-2003, but rebounds from 2004-2007. The comparisons of U.S. to Non-U.S. output of hESC articles, RNAi articles, and nearest neighbor articles are consistent with a general phenomenon in which the global concentration of science is declining, as countries other than the United States increase their relative investments in scientific research. The pattern is different, however, in (4-3), in which the control sample consists of Other Stem Cell root articles. hESC research is neither especially “hot” nor especially “cold” in comparison to OSC research and the policy shift does not appear to affect one to a significantly greater degree than the other, with the exception that U.S.-hESC output declines in 2002 and 2003. Although we believe that the comparison of hESC and OSC output is illustrative, we do not believe

that OSC forms a suitable control sample for hESC research, as it is relatively easy for researchers to substitute from hESC research into OSC research (and vice versa).

V.4. Mechanisms of Response and Sources of the Rebound

Our analyses thus far have focused on identifying the *extent* of U.S.-based hESC research during various U.S. policy environments. In Figures 4 and 5, we examine changes in the *nature* of U.S.-based hESC research, highlighting mechanisms that may partially explain the relative rebound in U.S. hESC output in the latter years of our sample. Specifically, Figure 4 investigates the relative impact of the policy shift on elite U.S. universities relative to other institutions, while Figure 5 investigates the impact of the policy shift on various types of research collaborations. Figure 4 compares the impact of the policy shock across three types of institutions: (a) elite U.S. universities, (b) all other U.S. institutions, and (c) institutions outside the United States. In these models, we have "triple-stacked" the data; coefficients on hESC*U.S.*Institution Type indicate the boost or decrement to the production of hESC papers by institutions of various types (Top-25 and Not-Top-25) in the U.S. in to the policy shift, each relative to citations from articles with Non-U.S. citing authors.

The (unreported) average treatment effects imply a relative decline of approximately 33 percent of output in the Top-25 U.S. universities and 46 percent among the non-elite universities. The year-by-year effects, however, imply that hESC output by the elite U.S. universities declined in 2002 and 2003, but recovered nearly completely thereafter. By contrast, hESC research output by U.S.-based reprint authors in other institutions experience a relative decline beginning in 2001, recovering somewhat, though not completely (either in magnitude or statistical significance) by the end of the study period.

These results are consistent with interview-based evidence we have assembled, which suggests that the constraints applied by the Bush Administration policy were more likely to be binding for those institutions for which federal funding was a relatively more important source of funding, whereas those institutions that found it easy to obtain private funding were less negatively impacted by the policy. Consider the Harvard Stem Cell Institute (HSCI) as an example of the type of research possible at elite institutions without federal funds. Harvard University established the HSCI in 2004, incorporating researchers and practitioners from Harvard's various colleges and affiliated hospitals. In 2005, the HSCI had begun to disperse nearly $1.8 million in research grants (HSCI, 2005); by 2006, the HSCI had received approximately $50 million in donor support (HSCI, 2006); and in 2009, the Institute dispersed $15.6 million in research funds, including some funds included in a five-year, $25 million targeted research grant from GlaxoSmithKline (HSCI, 2009). HSCI research on human embryonic stem cells is rigorously kept physically and organizationally

separate from research supported by federal funds (Dreifus, 2006) and several Harvard faculty maintain distinct laboratories for their hESC work. It does appear to have taken a few years, however, for elite scholars and institutions, like Harvard, to attract substantial private research funds and to have generated the infrastructure necessary to ensure compliance with federal policies and begin to produce frontier research. Moreover, the extent to which such substitution is possible may be limited both to a select set of institutions and to particular scientific fields, like hESC research, which are relatively small in comparison to other fields.

We examined an additional mechanism by which U.S.-based hESC researchers may have responded to the policy shock in Figure 5. Specifically, Figure 5 compares follow-on research across three collaboration-location types: (a) papers with only U.S. authors, (b) papers with U.S. and Non-U.S. authors (i.e., those with international collaboration), and (c) papers with no U.S. authors. The results suggest that the output of papers with only U.S.-based authors declined more significantly than those of the other types following the policy shock. The relative output of hESC papers involving collaboration between U.S. and non-U.S. authors declined significantly in the year of the policy shock (2001) but were unaffected by the policy shock thereafter. This is a particularly interesting mechanism of response by the U.S. hESC community, which suggests that international scientific networks may help mitigate the impact of any one nation's policies. (The result is also consistent with the expectations of Garud and Rappa (1995) who propose that collaborations will play a role in shaping researcher commitment to a particular research area.) Taken together, these results are consistent with an explanation in which researchers (possibly those at more resource-constrained institutions) in the U.S. collaborate with scientists outside the United States who may have less complex and better access to resources.

VI. Discussion

This paper presents a detailed analysis of the impact of the 2001 U.S. Human Embryonic Stem Cell policy on the relative output of U.S. research building on seminal discoveries in hESC research. Our results suggest that U.S. hESC research output experienced a statistically significant and quantitatively meaningful decline relative to hESC research output in the rest of the world, when comparing to the trend anticipated by the relative U.S. share of RNAi research output. The relative decline in U.S. hESC output is most severe in 2001, the year in which the uncertainty about U.S. federal funding policy was resolved.

The extent to which the observed patterns of US activity in stem cell research output are a direct consequence of the targeted funding policy *per se* or whether other unobservable factors

affected U.S. hESC research beginning in 2001 is open to interpretation. It is possible that *uncertainty* surrounding the 2001 policy, rather than the policy itself, may have discouraged research efforts during the early 2000s. The data make clear that the hESC policy environment (i.e., both the uncertainty and the policy), strongly influenced U.S. hESC research output relative to what would have been expected given the experience of RNAi. While our results are consistent with the observation that scientific and technical capabilities are increasingly globalized (Furman and Hayes, 2004), our analysis suggests that U.S. policies can accelerate the globalization of science if they employ funding targets that are inconsistent with the preferences of research scientists or place constraints on research autonomy within the United States.

Considering the limited support for U.S. hESC research and the much more explicitly supportive policy environments of a number of other countries, including South Korea, Singapore, and Israel, perhaps the biggest surprise in our results is the long-term robustness of the United States hESC research community to system perturbations. We thus interpret our overall findings as consistent with a picture in which the national-level institutions that support scientific competitiveness and scientific autonomy are robust and relatively enduring. Overall, once scientists' willingness to find alternative mechanisms to pursue their chosen paths is taken into account, the impact of the Bush Administration targeting may be of second-order importance relative to issues such as the extent of overall funding.[20]

By considering how the research trajectories of the U.S. scientific community responded to the complex policy context for hESC research in the 2001 (and earlier) period, our paper provides the first causal assessment of this highly controversial public policy (reversed in 2009 by order of the Obama Administration, subject to continued legal deliberation). Our evaluation of this changing policy environment addresses and clarifies a series of debates among policy-makers, including those who consider science policy in the context of national competitiveness (Marburger, 2005; Jaffe, 2006). Specifically, it provides insight into the degree to which national funding agencies (as well as those operating at the state level or in the private sector) can shape overall participation in a particular research trajectory, alter the composition of researchers in a particular research area, or affect the nature of collaboration in a field. Our results suggest that scientists exhibit considerable agency with regards to maintaining their research direction, thus implicitly raising questions as to the relative power of funders to shape the national research portfolio (Gans and Murray 2011).

A further contribution of our paper is to exemplify recent developments in science policy evaluation, often referred to as the "Science of Science Innovation Policy." By developing this paper

[20] We are grateful to Paula Stephan for pointing this out in careful Discussant commentary on this paper.

as both an empirical exercise in its own right and a primer on the citation-based approach to policy evaluation, we hope to have an impact on a range of scholars in the science policy community. Our analysis illustrates the benefits of the citation-based approach, while highlighting the difficulties associated with science policy evaluation, even using these advanced econometric techniques. At the core of the challenges is the ability to define a clear policy intervention, assess its likely impact on the scientific community, find appropriate measures of scientific activities in the pre- and post-period, and define a clear counterfactual. The citation-based approach provides a clear framework within which to undertake such as analysis. However, it is also limited in that is focuses specifically on having a precise understanding of the nature of the policy intervention and on examining the impact of the intervention on research trajectories. As we have shown in our earlier work (e.g., Murray and Stern, 2007; Furman and Stern, 2011; Furman, Jensen, and Murray, 2012), this approach is most powerfully applied to institutional and policy changes that are tightly coupled to specific research paths and research areas. Broadening to the more complex hESC policy is at once an opportunity and a challenge. The messy hESC situation imposes limits on the precision with which we can interpret the results of our analysis. Nonetheless, the approach provides complementary insights to the work of others in the hESC area who have focused on tracing out the activities of specific scientists, particular research areas or key research materials.

Lastly, and most broadly, our analysis of the stem cell policy environment illuminates some critical issues in the microeconomic and institutional foundations of scientific progress (Mokyr, 2002). Although "Open Science" is widely recognized to play a fundamental role in the production of fundamental knowledge (Merton, 1973; Dasgupta and David, 1994; David 1998, 2001, 2008; Stephan, 1996), few formal analyses support our understanding of the impact of specific and often subtle policies and practices on scientific progress. Beyond simple determination of the level of participation, our analysis suggests that policies may have important, though unintended, *distributional* consequences: For example, policies may have an impact on national scientific advantage by changing the degree to which researchers engage in a particular trajectory relative to the rest of the world. Our findings also remind us that scientific progress takes place in an increasingly globalized context, with scientists (particularly at elite universities) strategically acting to select international scientific collaborators from where policies may be less restrictive. This underscores the unintended consequences of implementing national policies in the context of the global scientific community.

REFERENCES

Adams, J. D. and Z. Griliches (1998) "Research Productivity in a System of Universities," *Annals of INSEE,* 49, 127-162.

Aghion, P. and Howitt, P. (1992) "A model of growth through creative destruction," *Econometrica,* 60, 323-351.

Aghion, P., Dewatripont, M., Kolev, J., Murray, F. and Stern, S. (2010) "The Public and Private Sectors in the Process of Innovation: Theory and Evidence from the Mouse Genetics Revolution," *American Economic Review: Papers and Proceedings*, 100, 153-158.

Aghion, P., M. Dewatripont, and J.C. Stein (2008) "Academic Freedom, Private-Sector Focus, and the Process of Innovation," *RAND Journal of Economics*, 39, 617 – 635.

Aghion, P., C. Harris, P. Howitt and J. Vickers (2001) "Competition, Imitation and Growth with Step-by-step Innovation", *Review of Economic Studies,* 68, 467-492.

Allison, P. D. and R. P. Waterman (2002) "Fixed effects negative binomial regression models." Ross Stolzenberg (ed.), *Sociological Methodology*. Boston: Basil Blackwell.

Anand, G. and A. Regalado (2002) "Slow Road For A Stem-Cell Researcher," *Wall Street Journal*, April 10, 2002, Wednesday, Section B; Page 1, Column 2.

Angrist, J. and A. Krueger (2001) "Instrumental Variables and the Search for Identification: From Supply and Demand to Natural Experiments," *Journal of Economic Perspectives,* 15, 69-85.

Azoulay, P., Waverly, D., and Stuart, T. (2007) "The Impact of Academic Patenting on (Public) Research Output," *Journal of Economic Behavior and Organization* 63, 599-623.

Azoulay, P., J. Graff Zivin, and G. Manso (2011) "Incentives and Creativity: Evidence from the Academic Life Sciences," *RAND Journal of Economics*, 42, 527-554.

Azoulay P., J. Graff Zivin & J. Wang (2010) "Superstar Extinction," *Quarterly Journal of Economics*, 125, 549-589.

Balakrishna, A. (2007) "Stem Cell Institute Aids Cooperation," *Harvard Crimson,* (June 06, 2007)

Beardsley, S. (2005) "A world of approaches to stem cell research," *Scientific American*, June 27, 2005.

Clark, J. (2001) "Squandering our technological future," *New York Times*, August 31, 2001.

Dasgupta, P. and P.A. David (1994) "Towards a New Economics of Science," *Research Policy* 23, 487-521.

David, P.A. (1998) "Common Agency Contracting and the Emergence of 'Open Science' Institutions," *American Economic Review*, 88, 15-21.

David, P.A. (2001) "Patronage, Reputation, and Common Agency Contracting in the Scientific Revolution: From Keeping 'Nature's Secrets' to the Institutionalization of 'Open Science,'" SIEPR Working Paper, Stanford University.

David, P.A. (2008) "The Historical Origins of 'Open Science': An Essay on Patronage, Reputation and Common Agency Contracting in the Scientific Revolution", *Capitalism and Society,* 3, Article 5.

De Solla Price, D. (1963) *Little Science, Big Science*, New York: Columbia University Press.

De Solla Price, D. (1965) "Networks of Scientific Papers", in *Science* 149, 510-515.

De Solla Price, D. (1967) "Nations can Publish or Perish", in *International Science and Technology* 70, 84-90.

Ding, W., F. Murray & T. Stuart (2006) "An Empirical Study of Gender Differences in Patenting among Academic Life Scientists." *Science*, 313, 665-667.

Dreifus, C. (2006) "At Harvard's Stem Cell Center, the Barriers Run Deep and Wide," *New York Times*, January 24, 2006.

Evans, M. and Kaufman, M. (1981) "Establishment in culture of pluripotent cells from mouse embryos," *Nature*, 292, 154–156.

Fletcher, L. (2001) "US stem cell policy comes under fire," *Nature Biotechnology* 19, 893 – 894.

Fuchs, E. (2009) "Cloning DARPA Successfully," *Issues in Science and Technology*. XXVI. (1)

Fuchs, E. (2010) "DARPA Does Moore's Law: The case of DARPA and Optoelectronic Interconnects," in F. Block and W. Keller, *State of Innovation: The U.S. Government's Role in Technology Development*, Paradigm Publishers.

Furman, J.L. and F. Murray (2008) "Autonomy & the Fragile Structure of the Scientific Community: Examining U.S. Stem Cell Policy," presentation to the American Sociological Association, 2008.

Furman, J.L, and Stern, S (2011) "Climbing Atop the Shoulders of Giants: The Impact of Institutions on Cumulative Research," *American Economic Review* 101, 1933-63.

Furman, J.L., K. Jensen & F. Murray (2012) "Governing Knowledge Production in the Scientific Community: Quantifying the Impact of Retractions". *Research Policy*, forthcoming.

Furman, J.L. and R. Hayes (2004) "Catching Up or Standing Still? National Innovative Productivity among 'Follower' Nations, 1978-1999, *Research Policy* 33, 1329-1354.

Gans J. and F. Murray (2011) "Public Funding Restrictions and their Impact on the Rate and Direction of Scientific Research," working paper, *Rate and Direction of inventive Activity, Conference in Honor of Kenneth Arrow and Richard Nelson.*

Garfield, E. (1955) "Citation indexes for science – New dimension in documentation through association of ideas," *Science*, 122, 108–111.

Garfield, E. (1979) "Is citation analysis a legitimate evaluation tool?" *Scientometrics*, 1(4), 359-375.

Garud, R. and M.A. Rappa (1995) "On the persistence of researchers in technological development," *Industrial and Corporate Change*, 15, 811-846.

Griliches, Z. (1990) "Patent Statistics as Economic Indicators: A Survey," *Journal of Economic Literature*, 92, 630-653.

Griliches, Z. (1998) *R&D and Productivity: The Econometric Evidence*, University of Chicago Press.

Hall, B., J. Mairesse, and L. Turner (2007) "Identifying Age, Cohort and Period Effects in Scientific Research Productivity: Discussion and Illustration Using Simulated and Actual Data on French Physicists," *Economics of Innovation and New Technology*, 16,159-177.

Harvard Stem Cell Institute (2005) *Harvard Stem Cell Institute Annual Report*, Harvard University: Cambridge, MA.

Harvard Stem Cell Institute (2006) *Harvard Stem Cell Institute Annual Report*, Harvard University: Cambridge, MA.

Harvard Stem Cell Institute (2009) *Harvard Stem Cell Institute Annual Report*, Harvard University: Cambridge, MA.

Hausman, J., B. Hall, and Z. Griliches (1984) Econometric Models for Count Data with an Application to the Patents R&D Relationship," *Econometrica*, 52, 909-938.

Hughes, R.J. (2006) "Stem-Cell Funding's Private Side," *Wall Street Journal*, July 28, 2006.

Imbens, G.W., and J.M. Wooldridge (2009) "Recent Developments in the Econometrics of Program Evaluation," *Journal of Economic Literature*, 47, 5–86.

Jaffe, A. (2006) "The 'Science of Science Policy': Reflections on the Important Questions and the Challenges They Present," Keynote Address at the NSF Workshop on Advancing Measures of Innovation: Knowledge Flows, Business Metrics, and Measurement Strategies.

Kuhn, T. S. (1962) *The Structure of Scientific Revolutions.* Chicago: University of Chicago Press.

Lane, J. (2009) "Assessing the Impact of Science Funding," *Science*, 324,1273-1275.

Lane, J. and S. Bertuzzi (2011) "Measuring the results from science investments," *Science*, 331, 678-680.

Levin, S. and P. Stephan (1991) "Research Productivity Over the Life Cycle: Evidence for Academic Scientists," *American Economic Review*, 81, 114-132.

Levine, A. (2005) "Geographic trends in human embryonic stem cell research," *Politics and the Life Sciences*, 23, 40-5.

Levine, A. (2008) "Identifying under- and overperforming countries in research related to human embryonic stem cells." *Cell Stem Cell*, 2,521-24.

Lindsey, D. (1989) "Using citation counts as a measure of quality in science: Measuring what's measurable rather than what's valid," *Scientometrics*, 15, 189-203.

Löser, P, J. Schirm, A. Guhr, A.W. Wobus, A. Kurtz (2009) "Human Embryonic Stem Cell Lines and Their Use in International Research," *Stem Cells*, 28, 240-246.

MacKinnon, James G. (2002) "Bootstrap inference in econometrics," *Canadian Journal of Economics*, 35, 615-645.

Marburger, J. (2005) "Wanted: Better Benchmarks," *Science*, 308, 1087.

Martin, G. (1981) "Isolation of a pluripotent cell line from early mouse embryos cultured in medium conditioned by teratocarcinoma stem cells," *Proceedings of the National Academy of Sciences, USA*, 78, 7634–7638.

Martin M.J., A. Muotri, F. Gage, and A. Varki (2005) "Human embryonic stem cells express an immunogenic nonhuman sialic acid," *Nature Medicine*, 11, 228-232.

McCormick, J, Owen-Smith, J and Scott, C. (2009) "Distribution of human embryonic stem cell lines: Who, when and where." *Cell Stem Cell* 4, 107-10.

McGinley, L. and A. Regalado (2002) "Cloning Clamor Grows," *Wall Street Journal*, pp. Section B; Page 1, Column 2 A1, B1,3 (April 10, 2002).

Merton, R.K. (1973) *The Sociology of Science: Theoretical and Empirical Investigations.* Edited by Norman Storer. Chicago: University of Chicago Press.

Mokyr, J. (2002) *The Gifts of Athena.* Princeton, NJ: Princeton University Press.

Murray, F (2010) "The Oncomouse that Roared: Hybrid Exchange Strategies as a Source of Productive Tension at the Boundary of Overlapping Institutions," *American Journal of Sociology* 116, 341-388.

Murray, F. (2007) "The Stem-Cell Market – Patents andd the Pursuit of Scientific Progress," *The New England Journal of Medicine*, 356, 2341-2343.

Murray, F. and S. Stern (2007) "Do formal Intellectual Property Rights Hinder the Free Flow of Scientific Knowledge? An Empirical Test of the anti-Commons Hypothesis," *Journal of Economic Behavior and Organization* 63, 648-687.

Murray, F., P. Aghion, M. Dewatripont, J. Kolev and S. Stern (2011) "Of Mice and Academics: The Role of Openness in Science," MIT Sloan Working Paper.

NIH (1999) "Fact Sheet on Stem Cell Research," Office of the NIH Director, Bethesda, MD.

NIH (2001) *Stem Cells: Scientific Progress and Future Research Directions*. U.S. Department of Health and Human Services.

NIH (2011) *Estimates of Funding for Various Research, Condition, and Disease Categories (RCDC)*, http://report.nih.gov/rcdc/categories/default.aspx, accessed by J. Furman, November 10, 2011.

Owen-Smith, J. and J. McCormick (2006) "An International Gap in Human Embryonic Stem Cell Research," *Nature Biotechnology*, 24, 391-392.

Rabb, H. (1999) "Memorandum from Harriet Rabb to Harold Varmus," January 15, 1999.

Rewerski, P. (2007) "The Need for a New U.S. Stem Cell Research Policy: A Comparative Look at International Stem Cell Research Laws," *Journal of Law, Technology, and Policy*, 415-432.

Romer, P.M. (1990) "Endogenous Technological Change," *Journal of Political Economy*, 98, S71-S102.

Schubert, A. and T. Braun (1993) "Reference standards for citation based assessments," *Scientometrics*, 26, 21-35.

Scotchmer, S (1991) "Standing on the Shoulders of Giants: Cumulative Research and the Patent Law," *Journal of Economic Perspectives* 5, 29-41.

Scott, C.T., J. McCormick, & J. Owen-Smith (2009) "And then there were two: Use of hESC lines," *Nature Biotechnology* 27, 696-697.

Scott, C.T., J. McCormick, M. DeRouen, & J. Owen-Smith (2010) "Federal policy and the use of pluripotent stem cells," *Nature Methods*, 11: 866-867.

Scott, C.T., J. McCormick, M. DeRouen, & J. Owen-Smith (2011) "Democracy Derived: New Trajectories in Pluripotent Stem Cell Research," *Cell*, 145, 820-826.

Smaglik, P. (2001) "Stem-cell seduction," *Nature* 412, 3-3.

Stephan. P. (1996) "The Economics of Science," *Journal of Economic Literature*, 34, 1199-1235.

Stokes, D.E. (1997) *Pasteur's Quadrant Basic Science and Technological Innovation*, Washington, DC: Brookings Institution Press.

Thomson, J.A. Itskovitz-Eldor, J., Shapiro, S.S., et al. (1998) "Embryonic stem cell lines derived from human blastocysts," *Science* 282, 1145-1147.

Vestal, C. (2007) "Embryonic stem cell research divides states," Stateline.org, Thursday, June 21, 2007; accessed June 30, 2010.

Vestal, C. (2009) "States applaud new stem-cell funding," Stateline.org, Wednesday, March 11, 2009; accessed June 30, 2010.

Vogel, G. (2001) "Bush squeezes between lines on stem cells," *Science*, 293, 1242 – 1245.

Walters, L. (2004) “Human Embryonic Stem Cell Research: An Intercultural Perspective,” *Kennedy Institute of Ethics Journal*, 14, 3–38.

Wertz, D.C. (2002) “Embryo and stem cell research in the United States: History and politics,” *Gene Therapy*, 9, 674–678.

Williams, H. (2010) “Intellectual property rights and innovation: Evidence from the human genome,” NBER Working Paper 16213.

Table 1 – Descriptive Statistics – Root Articles

Panel A: Characteristics of Articles, entire sample*

Variable *(n=719)*	**Mean**	**Std. Dev.**	**Min**	**Max**
Publication Year	1997.45	3.94	1976	2002
Reprint Author based in US	0.61	0.49	0	1
Reprint Author based in University	0.68	0.47	0	1
Reprint Author at Top 25 US University	0.24	0.43	0	1
Paper with Only US author(s)	0.47	0.50	0	1
Paper with Any US author(s)	0.59	0.49	0	1
Paper with both US & Non-US authors	0.13	0.34	0	1
Paper with No US authors	0.39	0.49	0	1

Panel B: Characteristics of Articles, by article type*

	hESC *(n=16)*		**RNAi *(n=51)***		**Nearest Neighbors *(n=559)***		**Other Stem Cells *(n=93)***	
Variable	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
Publication Year	1995.19	6.01	1999.84	2.61	1997.36	3.81	1997.11	4.41
Reprint Author based in US	0.33	0.49	0.76	0.43	0.61	0.49	0.56	0.50
Reprint Author based in Uni	0.87	0.35	0.56	0.50	0.68	0.47	0.69	0.47
Reprint Author at Top 25 US Uni	0.27	0.46	0.20	0.40	0.25	0.43	0.19	0.39
Paper with Only US author(s)	0.19	0.40	0.59	0.50	0.47	0.50	0.43	0.50
Paper with Any US author(s)	0.31	0.48	0.71	0.46	0.59	0.49	0.52	0.50
Paper with both US & Non-US authors	0.14	0.36	0.13	0.34	0.14	0.35	0.09	0.29
Paper with No US authors	0.67	0.49	0.24	0.43	0.39	0.49	0.44	0.50
Reprint Author blank	0.06	0.25	0.12	0.33	0.20	0.40	0.25	0.43
Address/Institution field blank	0.13	0.34	0.08	0.27	0.09	0.29	0.05	0.23

** The percentage of US + Non-US authors does not sum to 1, as Reprint Author data are blank (not specified or not recorded) for 142 of the 719 root articles.*

Panel C: Counts of Reprint Author by Country of Origin (for top nine countries of origin)^

hESC		**RNAi**		**NN***		**OSC***	
RP Country	*# papers*	*RP Country*	*# papers*	*RP Country*	*# papers*	*RP Country*	*# papers*
USA	5	USA	34	USA	271	USA	39
Israel	4	Germany	4	England	27	Japan	5
Canada	2	Netherlands	3	Germany	26	England	4
England	2	Italy	1	Australia	16	Germany	7
		Japan	1	France	15	Australia	2
		Norway	1	Japan	11	Austria	2
				Italy	10	Canada	2
				Canada	9	France	2
				France	8	Italy	2

^ Note that Reprint Author addresses are not available for all papers.

** Table omits countries with fewer than two RP authors on OSC papers and fewer than eight RP authors on NN root articles.*

Table 2 – Descriptive Statistics – Characteristics of Annual Citations Received by Root Articles

Panel A: Characteristics of Annual Citations Received (citation-year observations), entire sample*

Variable	Mean	Std. Dev.	Min	Max
Citing Year	2002.54	3.06	1996	2007
Annual Citations	17.26	39.36	0	698
Annual citations received from articles with…				
Reprint Author based in US	7.33	16.66	0	305
Reprint Author based outside US	9.93	23.65	0	516
Reprint Author based at University	11.12	25.65	0	487
Reprint Author based at Top 25 US University	2.39	5.55	0	95
Paper include Any US author(s)	8.30	18.46	0	344
Paper include Only US author(s)	6.50	14.71	0	270
Paper with both US & Non-US authors	1.81	4.13	0	74
Paper with no US authors	8.95	21.74	0	477
Papers in which RP author missing	0.24	0.43	0	1
Papers in which Address data (C1) missing	0.11	0.31	0	1

Panel B: Characteristics of Annual Citations Received, by article type*

	hESC *(n=159)*		***RNAi*** *(n=403)*		***Nearest Neighbors*** *(n=5382)*		***Other Stem Cells*** *(n=901)*	
Variable	**Mean**	**Std. Dev.**	**Mean**	**Std. Dev.**	**Mean**	**Std. Dev.**	**Mean**	**Std. Dev.**
Citing Year	2002.32	3.19	2003.33	2.68	2002.50	3.07	2002.47	3.09
Annual Citations	32.03	55.05	67.95	94.61	8.69	16.01	43.18	58.45
Annual Citations received from articles with…								
Reprint Author based in US	11.06	19.43	31.53	41.80	3.84	7.44	16.70	22.50
Reprint Author based outside US	20.97	36.19	36.42	54.31	4.85	9.35	26.48	37.66
Reprint Author based at University	20.26	35.15	41.39	59.42	5.74	11.01	28.09	39.96
Reprint Author based at Top 25 US University	4.14	7.66	10.18	13.64	1.27	2.75	5.25	7.22
Paper include Any US author(s)	12.83	22.02	34.49	46.24	4.40	8.26	19.13	24.99
Paper include Only US author(s)	9.60	16.77	28.12	36.63	3.36	6.53	15.00	19.98
Paper with both US & Non-US authors	3.23	5.79	6.37	10.16	1.04	2.11	4.13	5.70
Paper with no US authors	19.19	33.44	33.47	49.67	4.29	8.48	24.05	34.95
Papers in which RP author missing	0.08	0.26	0.16	0.37	0.24	0.43	0.30	0.46
Papers in which Address data (C1) missing	0.15	0.36	0.10	0.31	0.11	0.31	0.06	0.24

Table 2 (continued) – Descriptive Statistics – Characteristics of Citing Articles

Panel C: Citing Articles by RP Author country of origin (citing articles published in or before 2001)

hESC		RNAi		NN		OSC	
RP Country	*# papers*	*RP Country*	*# papers*	*RP Country*	*# papers*	*RP Country*	*# papers*
USA	442	USA	1,383	USA	7,266	USA	3,883
England	122	England	258	England	1,058	Germany	686
Germany	64	Germany	166	Germany	1,041	Japan	567
Australia	61	France	152	Japan	1,012	England	531
Japan	45	Japan	126	France	760	France	482
Israel	43	Netherlands	90	Canada	506	Canada	427
France	38	Austria	65	Italy	382	Italy	334
Canada	36	Australia	59	Switzerland	319	Australia	215
Netherlands	35	Italy	58	Netherlands	278	Sweden	132

Panel D: Citing Articles, RP Author country of origin (citing articles published in or after 2002)

hESC		RNAi		NN		OSC	
RP Country	*# papers*	*RP Country*	*# papers*	*RP Country*	*# papers*	*RP Country*	*# papers*
USA	1,604	USA	12,708	USA	16,083	USA	13,387
England	461	Japan	2,314	Japan	2,812	Japan	4,208
Israel	348	Germany	2,119	Germany	2,758	Germany	2,996
Japan	330	China	1,793	England	2,384	England	2,242
Germany	267	England	1,243	France	1,469	China	1,829
China	181	France	843	Canada	1,331	Italy	1,635
South Korea	173	Netherlands	665	Italy	1,251	France	1,319
Australia	168	Canada	605	China	983	Canada	1,206
Canada	155	Italy	538	Netherlands	806	South Korea	861
Singapore	117	South Korea	461	Spain	663	Australia	675

Table 3: Core Results – US hESC Output vs. Rest-of-World (1996-2007)

	CONDITIONAL FIXED EFFECTS NEG BINOMIAL, STACKED *[Incidence-Rate Ratios in brackets in top line]* *Estimated coefficients in 2nd line* *(Block bootstrapped SEs reported in parentheses)* ***DV = Cites with U.S. Reprint Author (or Not U.S. Reprint Author)***		
	(3-1)	**(3-2)**	**(3-3)**
HESC*2001	[1.081] 0.078 (0.203)	[1.190] 0.174 (0.298)	
HESC*(2002-2007)	[1.079] 0.076 (0.336)		
HESC*(2002-2003)		[1.073] 0.070 (0.323)	
HESC*(2004-2007)		[1.284] 0.250 (0.346)	
US*HESC*2001	**[0.488]** **-0.718** **(0.210)*****	**[0.489]** **-0.715** **(0.227)*****	
US*HESC*(2002-2007)	**[0.589]** **-0.530** **(0.108)*****		
US*HESC*(2002-2003)		**[0.413]** **-0.883** **(0.134)*****	
US*HESC*(2004-2007)		**[0.711]** **-0.341** **(0.111)*****	
HESC*1996			[1.332] 0.287 (0.831)
HESC*1997			[1.534] 0.428 (0.921)
HESC*1998			[1.507] 0.410 (0.891)
HESC*1999			[1.591] 0.464 (0.824)
HESC*2000			[0.905] -0.099 (0.765)
HESC*2001			[1.343] 0.295 (0.715)
HESC*2002			[1.144] 0.134 (0.747)
HESC*2003			[1.275] 0.243 (0.821)
HESC*2004			[1.287] 0.252 (0.815)
HESC*2005			[1.480] 0.392 (0.827)
HESC*2006			[1.653] 0.503 (0.842)
HESC*2007			[1.474] 0.388

			(0.885)
US*HESC*1996			[1.197] 0.180 (0.687)
US*HESC*1997			[1.057] 0.055 (3.016)
US*HESC*1998			[0.855] -0.157 (0.636)
US*HESC*1999			[0.635] -0.454 (0.259)*
US*HESC*2000			[0.891] -0.116 (0.217)
US*HESC*2001			**[0.491]** **-0.712** **(0.244)*****
US*HESC*2002			**[0.425]** **-0.856** **(0.218)*****
US*HESC*2003			**[0.402]** **-0.912** **(0.145)*****
US*HESC*2004			**[0.667]** **-0.405** **(0.158)****
US*HESC*2005			**[0.737]** **-0.305** **(0.116)*****
US*HESC*2006			**[0.659]** **-0.418** **(0.159)*****
US*HESC*2007			[0.793] -0.231 (0.211)
Article*Year Observations	1124	1124	1124
Number of articles	67	67	67
Log Likelihood	-3187.15	-3169.81	-3161.86

*Models include unreported constant, hESC*YearFEs, Stack-specific Year FEs, Stack FEs, Article Age FEs, and Article FEs*
Bootstrapped standard errors in parentheses
** significant at 10%; ** significant at 5%; *** significant at 1%*

Table 4: Exploring robustness of core results to an alternative measure of location and to alternative control samples

	CONDITIONAL FIXED EFFECTS NEG BINOMIAL, STACKED *[Incidence-Rate Ratios in brackets in top line]* *Estimated coefficients in 2nd line* *(Block bootstrapped SEs reported in parentheses)*		
	(4-1) ***Sample: hESC & RNAi articles*** ***DV = Cites with <u>any</u> US author (or No US author)***	**(4-2)** ***Sample: hESC & nearest neighbor articles*** ***DV = Cites with US Reprint Author (or Not US Reprint Author)***	**(4-3)** ***Sample: hESC & OSC articles*** ***DV = Cites with US Reprint Author (or Not US Reprint Author)***
HESC*2001	[1.250] 0.224 (0.212)	**[1.594]** **0.466** **(0.115)*****	[0.999] -0.001 (0.205)
HESC*(2002-2003)	[1.078] 0.076 (0.292)	**[1.918]** **0.651** **(0.163)*****	[0.853] -0.159 (0.239)
HESC*(2004-2007)	[1.333] 0.287 (0.351)	**[2.579]** **0.947** **(0.213)*****	[0.898] -0.107 (0.272)
US*HESC*2001	**[0.492]** **-0.709** **(0.155)*****	**[0.627]** **-0.467** **(0.210)****	[0.716] -0.335 (0.232)
US*HESC*(2002-2003)	**[0.477]** **-0.741** **(0.144)*****	**[0.540]** **-0.616** **(0.142)*****	**[0.641]** **-0.445** **(0.139)*****
US*HESC*(2004-2007)	**[0.731]** **-0.314** **(0.069)*****	**[0.775]** **-0.255** **(0.105)****	[1.050] 0.049 (0.107)
Article*Year Observations	1124	10964	2102
Number of articles	67	569	108
Log Likelihood	-3187.15	-17938.24	-5644.91

*Models include unreported constant, hESC*YearFEs, Stack-specific Year FEs, Stack FEs, Article Age FEs, and Article FEs*
Bootstrapped standard errors in parentheses
** significant at 10%; ** significant at 5%; *** significant at 1%*

Figure 1 – Citing articles by year

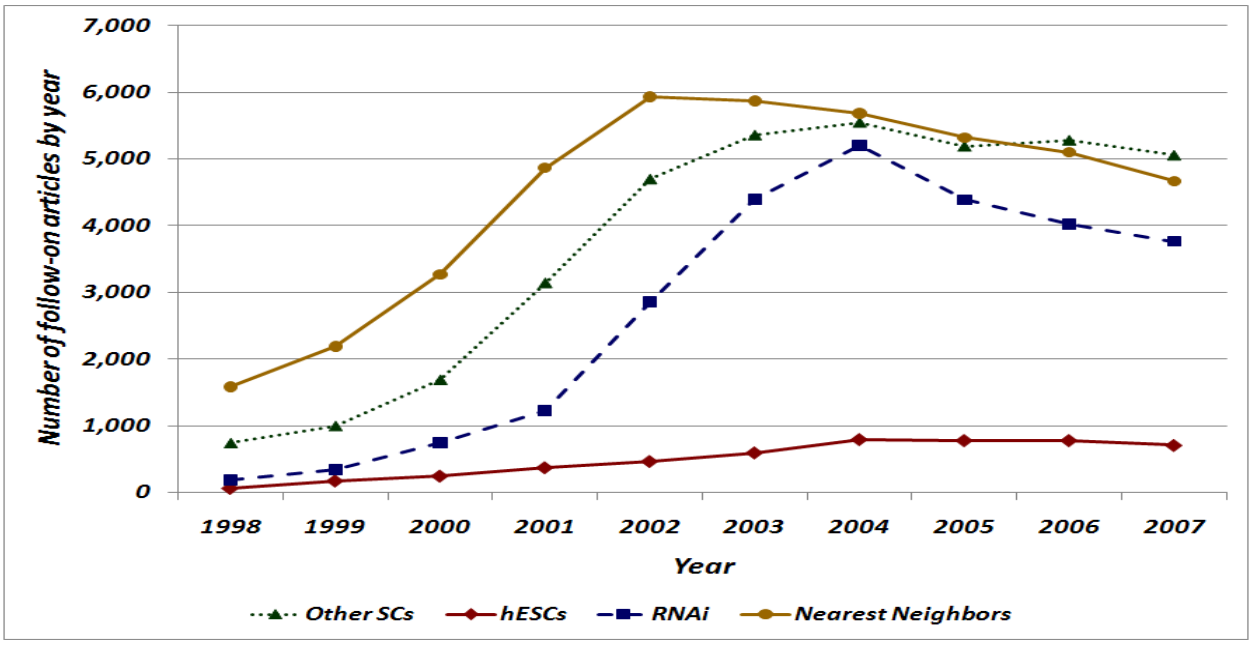


Figure 2 – Citing articles by publication type and year, U.S. vs. Non-U.S. Reprint Authors

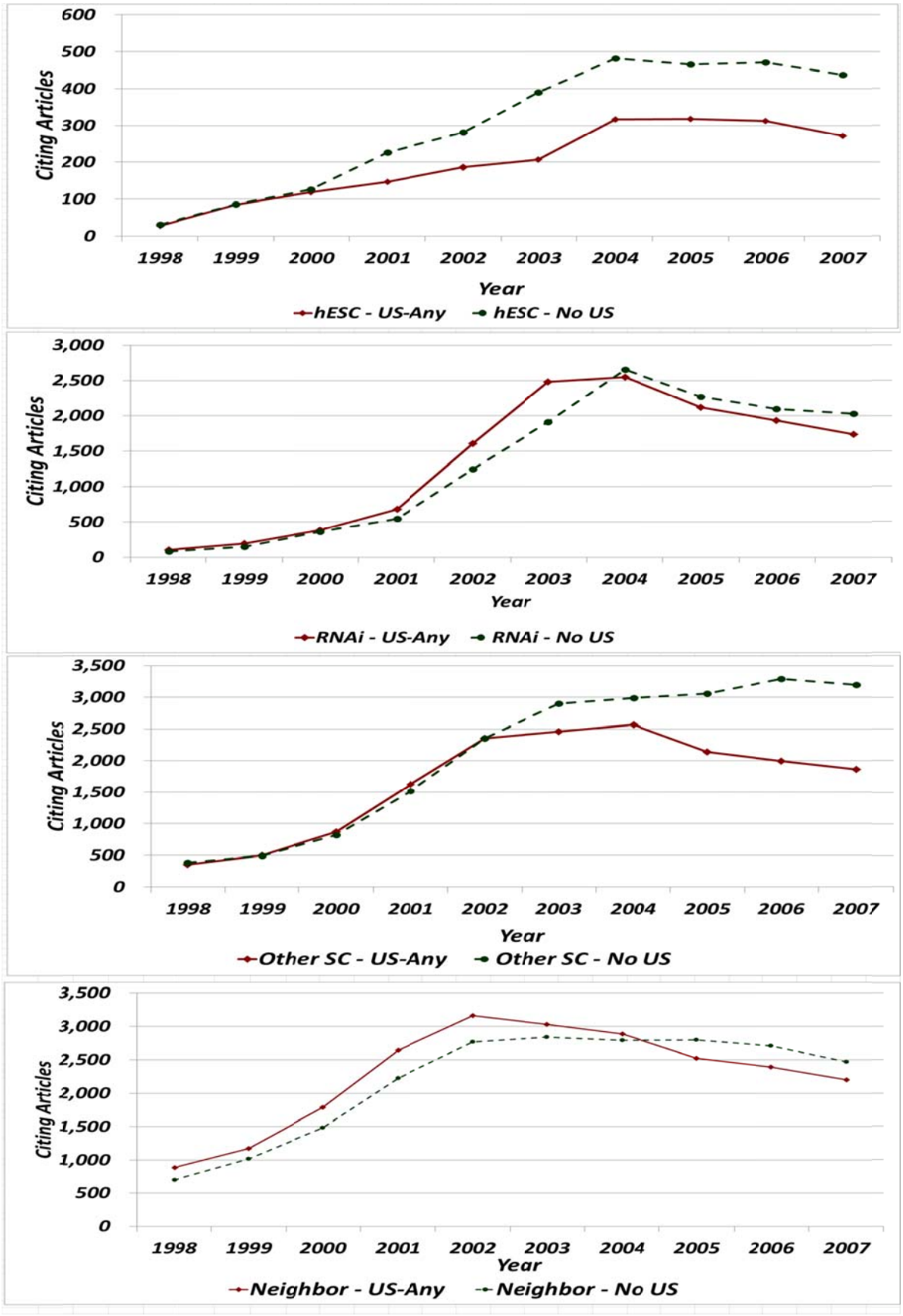


Figure 3 – U.S. hESC Output vs. Rest-of-World (1996-2007)

Figure 3 graphs the year-by-year incident rate ratios on hESC*US obtained in (3-3), including the upper and lower bounds on the 95% confidence intervals implied by the bootstrapped standard errors.

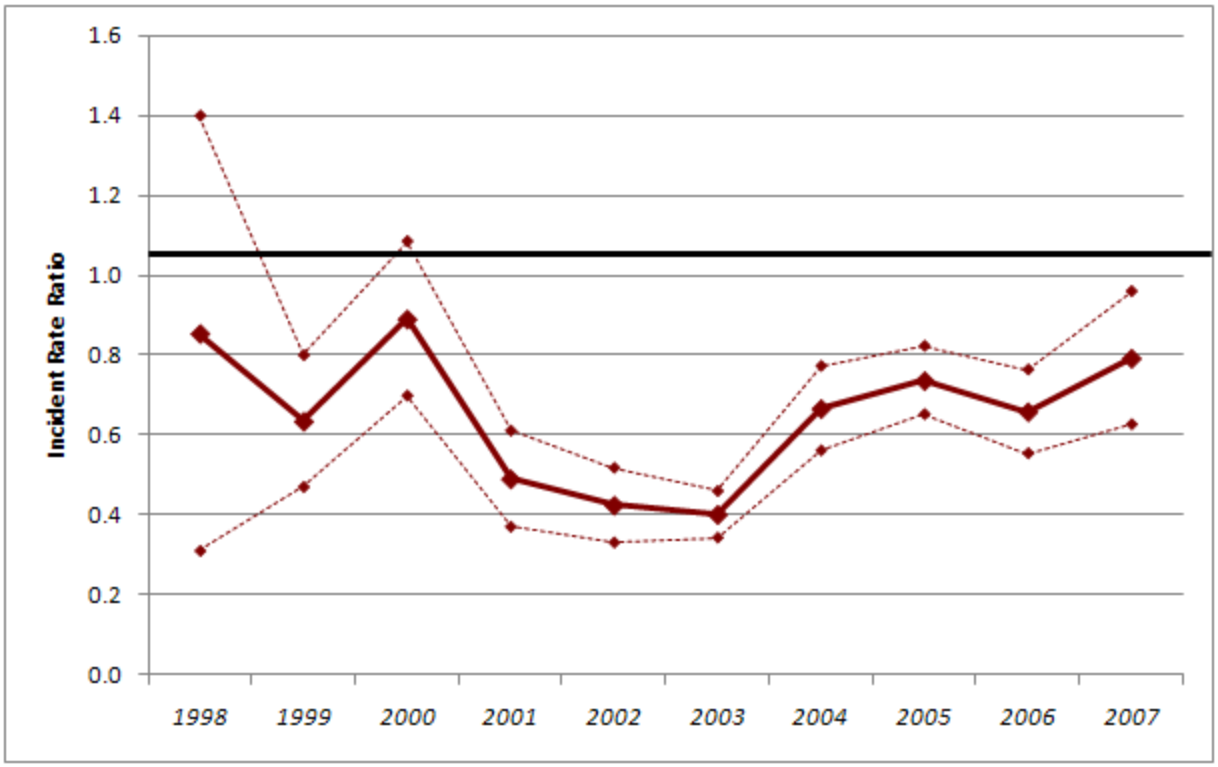


Figure 4 –hESC research output by Elite U.S. Universities & All Other U.S. Institutions

Figure 4 reports the results of conditional fixed negative binomial regressions with a structure similar to (3-3), which contains three stacks of the data, reflecting citations by papers with (a) Reprint Authors based at the Top 25 U.S. universities, (b) all other U.S. institutions, and (c) institutions outside the United States. Specifically, the figure reports incident rate ratios for year-specific variables reflecting hESC*(U.S. TOP 25 UNIVERSITIES) and hESC*(ALL OTHER U.S. INSTITUTIONS). The figure also reports confidence intervals based on bootstrapped standard errors. The regressions also include unreported constant, hESC*YearFEs, Stack-specific Year FEs, Stack FEs, Article Age FEs, and Article FEs.

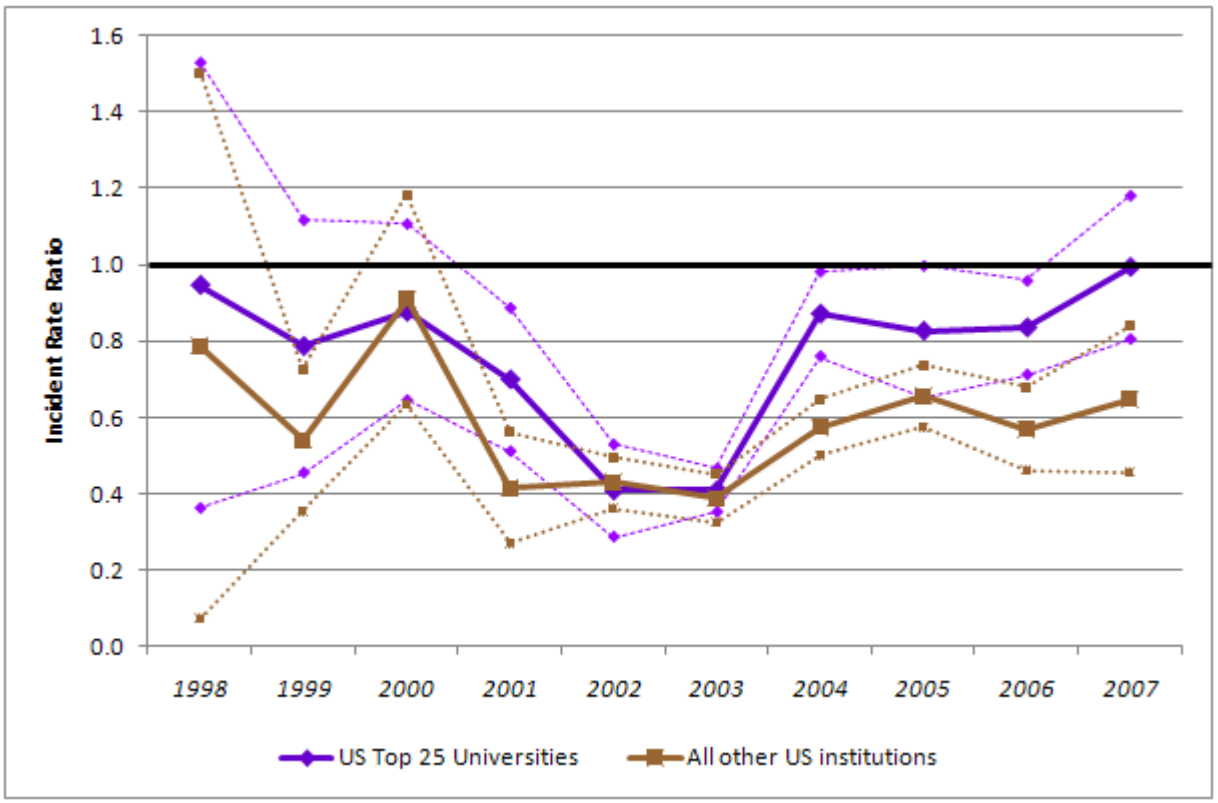
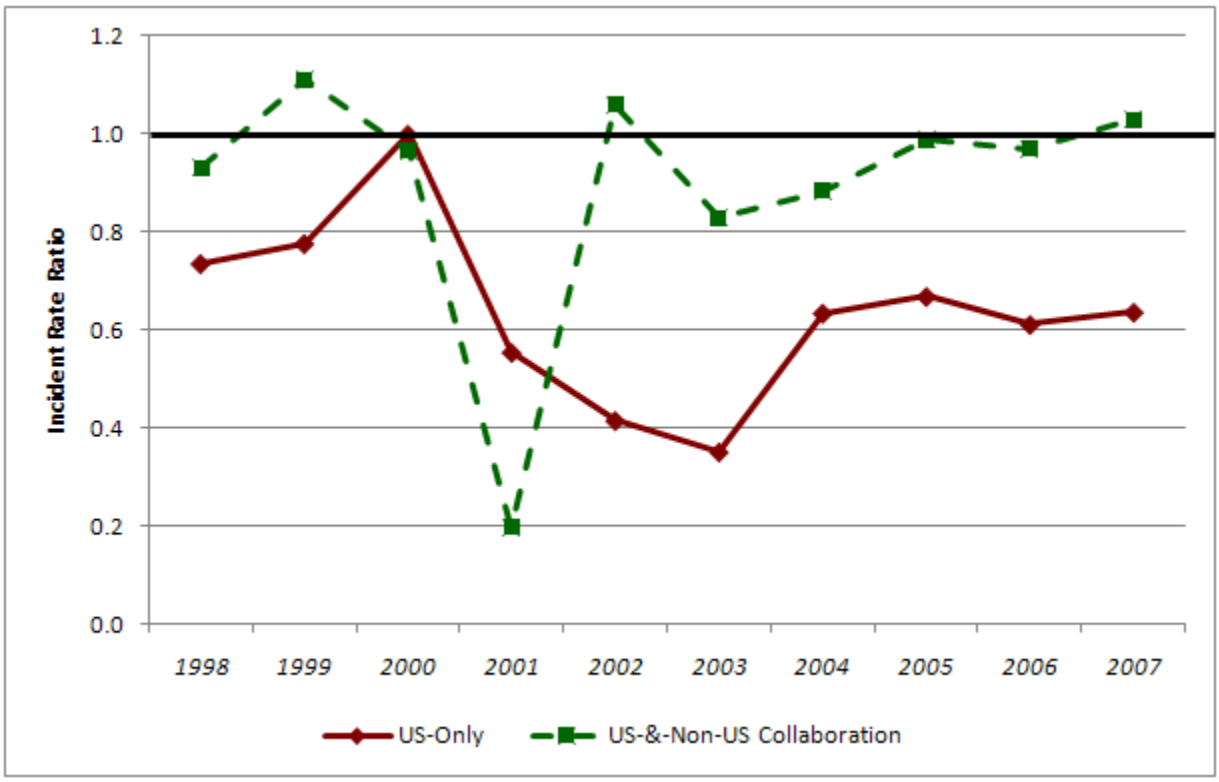


Figure 5 –Collaboration between U.S. & Non-U.S. authors

Figure 5 reports the results of conditional fixed negative binomial regressions of the type reported in (3-3), which contains three stacks of the data, which reflect citations by papers with papers with only U.S. authors, (b) papers with U.S. and Non-U.S. authors, and (c) papers with no U.S. authors. Specifically, the figure reports incident rate ratios for year-specific variables reflecting hESC*(US ONLY AUTHORS) and hESC*(US&NON_US COLLABORATION). The regressions also include unreported constant, hESC*YearFEs, Stack-specific Year FEs, Stack FEs, Article Age FEs, and Article FEs.



Appendix Table 1:
Citing Papers involving Collaborations between U.S. and other country, by sample and time period

hESC	2001 and prior	Post-2001
Germany	20	62
England	18	59
Japan	12	45
Canada	11	49
Italy	9	32
France	9	20
Israel	9	37
Sweden	8	20
Netherlands	7	24
Switzerland	3	6
Australia	3	33
Belgium	2	4
Denmark	2	1
Poland	1	5
Czech	1	4
China	0	23
South Korea	0	30
Singapore	0	21
Spain	0	7
Taiwan	0	11

RNAi	2001 and prior	Post-2001
France	45	196
Germany	32	309
England	31	195
Canada	29	192
Netherlands	29	110
Switzerland	29	113
Japan	28	307
Australia	15	66
South Korea	13	78
Singapore	8	35
Sweden	6	67
Israel	4	81
Taiwan	4	32
Italy	3	115
China	3	274
Czech	3	19
Poland	2	18
Spain	2	61
Denmark	2	28
Belgium	0	29

* Collaboration defined as paper that includes at least one U.S.-based address and at least one country other than the U.S. in the address field. Note that same paper may involve collaborations with multiple other countries.

Appendix Table 2 – Panel A:
hESC Citing Papers, Ranked by papers per Institutions with U.S.-based Reprint Authors

Institution	2001 & Prior
U WISCONSIN*	26
WISTAR INST ANAT & BIOL	18
U MICHIGAN*	19
U PENNSYLVANIA*	17
NINDS	14
JOHNS HOPKINS U*	11
THOMAS JEFFERSON U	11
HARVARD U*	9
MEM SLOAN KETTERING CANC CTR	9
U WASHINGTON*	9
WASHINGTON U*	9
NCI	8
U MINNESOTA*	8
GERON CORP	7
MICHIGAN STATE U	7
NINCDS	7
U FLORIDA	7
BAYLOR COLL MED	6
OREGON HLTH SCI U	6
ADV CELL TECHNOL	5
MASSACHUSETTS GEN HOSP	5
OREGON REG PRIMATE RES CTR	5
U ILLINOIS	5
U SOUTH FLORIDA	5

Institution	Post-2001
JOHNS HOPKINS U*	98
U WISCONSIN*	76
HARVARD U*	59
STANFORD U*	51
GERON CORP	47
NIA	43
U MINNESOTA*	41
MIT*	40
UC SAN DIEGO	32
WAKE FOREST U	32
UC SAN FRANCISCO*	28
U FLORIDA	24
U TEXAS	23
U SOUTHERN CALIFORNIA	22
U PITTSBURGH	21
CHILDRENS HOSP	20
COLUMBIA U*	18
U GEORGIA	18
U WASHINGTON*	18
BAYLOR COLL MED	16
U PENNSYLVANIA*	16
WHITEHEAD INST BIOMED RES	16
UC LOS ANGELES	15
WASHINGTON U*	15

* Indicates institution classified as Top 25 U.S. research university.

Appendix Table 2 – Panel B:
hESC Citing Papers, Ranked by papers per Institutions with Non-U.S.-based Reprint Authors

Institution	2001 & Prior
MONASH UNIV	42
UNIV CAMBRIDGE	23
NETHERLANDS INST DEV BIOL	18
UNIV SHEFFIELD	18
HEBREW UNIV JERUSALEM	16
UNIV EDINBURGH	11
UNIV OXFORD	11
TECHNION ISRAEL INST TECH	10
UNIV MIGUEL HERNANDEZ	10
CNR	9
OSAKA UNIV	9
UNIV BONN	9
WELLCOME CRC INST	8
HADASSAH UNIV HOSP	7
INST PLANT GENET & CROP PLANT RES	7
UNIV TORONTO	7
CHIBA UNIV	6
INST PASTEUR	6
KYOTO UNIV	6
MT SINAI HOSP	6
UNIV LIVERPOOL	6
INST CHILD HLTH	5
IPK GATERSLEBEN	5
MAX PLANCK INST IMMUNBIOL	5
MED UNIV LUBECK	5
PPL THERAPEUT	5
UNIV COLOGNE	5

Institution	Post-2001
HEBREW UNIV JERUSALEM	108
TECHNION ISRAEL INST TECHNOL	94
MONASH UNIV	83
UNIV SHEFFIELD	73
RAMBAM MED CTR	71
KYOTO UNIV	53
NATL UNIV SINGAPORE	51
SEOUL NATL UNIV	45
UNIV CAMBRIDGE	41
UNIV LONDON IMPERIAL COLL	41
UNIV TORONTO	41
UNIV EDINBURGH	34
JOHN P ROBARTS RES INST	33
UNIV DURHAM	31
UNIV NOTTINGHAM	30
CELLARTIS AB	27
NETHERLANDS INST DEV BIOL	27
ROSLIN INST	27
LUND UNIV	26
MIZMEDI HOSP	25
SHINSHU UNIV	21
UNIV COLOGNE	21
OSAKA UNIV	20
ROYAN INST	20
SHANGHAI JIAO TONG UNIV	20
TEL AVIV UNIV	20
UNIV BONN	20