Growing Stem Cells: The Impact of US Policy on the Geography and Organization of Scientific Discovery*

Jeffrey L. Furman
Boston University & NBER

Fiona Murray
MIT Sloan School of Management

Scott Stern
Northwestern University, MIT-Sloan, & NBER

March 1, 2010

[Preliminary]

* Acknowledgements: We thank Kevin Eggan, Amy Finkelstein, Julia Lane, Chris Liu, Megan MacGarvie, Matt Notowidigdo, Paula Stephan, and participants in numerous conferences and seminars for comments and suggestions. Devin Fensterheim, Pegi Fogertey, Anna Harrington, and Brianna Pesci provided excellent research assistance. All errors are our own. Financial support for this research was provided by the 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, Kellogg Graduate School of Management, 2001 Sheridan Road, Evanston, IL 60208 s-stern2@northwestern.edu.

Growing Stem Cells: The Impact of US Policy on the Geography and Organization of Scientific Discovery

Abstract

We investigate in this paper the impact of the August 2001 U.S. administration stem cell research policy on the extent and nature of scientific output. The 2001 policy constitutes an interesting policy experiment, as its impact on the field was profound and its particular form was unanticipated in advance. The policy enabled the first U.S. federal funding for human embryonic stem cell (hESC) research; however, it precluded the use of federal funds for all but a narrow (and scientifically unpromising) set of pre-existing stem cell lines. We evaluate the specific impact of this policy on (a) the extent of U.S. stem cell research relative to the international community and (b) the consequences of this policy for the composition of follow-on research in the United States. A particular challenge in this research is in identifying a counterfactual estimate of the production of human embryonic stem cell research that would have occurred had the policy shock not been implemented. To address this issue, we develop multiple control samples, including a particularly novel sample based on the production of research in RNA Interference (RNAi), a scientific breakthrough in a closely related field in cell biology that occurred in the same year as human embryonic stem cell research (1998) and that, like HESC research, also pioneered in the United States. Our estimates suggest that the production of human embryonic stem cell research in the United States was approximately 35-40 percent lower following the policy shift than it would have been in the absence of the shift. This decline appears to have been largely concentrated in the years 2001-2003. The impact of the policy is lessened, both in the sense of economic and statistical magnitude in the later years of the sample and publications by faculty in elite journals and those involving US collaborations with the rest of the international community do not experience a long-term decline. These results suggest that life sciences research in the U.S. is relatively robust in the face of such nuanced shocks and that such shocks engender a significant behavioral response on the part of U.S. scientists.

I. Introduction

In this paper we investigate the impact of the August 2001 U.S. Administration human embryonic stem cell (hESC) research policy on the extent and nature of hESC science in the United States relative to the rest of the world. The executive order issued by President George W. Bush in August 2001 clarified a period of policy uncertainty, which had existed since hESC research was enabled in 1998. Specifically, the Bush Administration policy provided the first federal funding for hESC research, although it also restricted that funding to support research employing a narrowly-defined set of cell lines that had been developed prior to policy's enactment. Thus, relative to an approach that did not involve funding restrictions, the Bush Administration policy limited scientists' ability to develop novel materials (with fewer contaminants or more genetic diversity) and pursue entirely new research directions. Policies regarding scientific research on early life materials have been historically controversial, and the Bush Administration policy regarding hESC stem cell research generated substantial debate within the U.S. scientific community, political sphere, media, and society. In 2001, research into the biological foundations of stem cells was described by some biologists as one of the most promising areas of scientific progress and the United States had been an early leader in this field. Critics of the administration policy feared that the severity of its restrictions would dramatically reduce U.S. investment in this area, reshaping global leadership in stem cell research and, perhaps, life sciences research more broadly and inhibiting the progress of knowledge in this area. Little systematic analysis, however, has examined these claims or evaluated the specific impact of U.S. human embryonic stem cell research (hESC) policy.

This paper exploits both the uncertainty in hESC policy and the surprising policy resolution in 2001 to assess the sensitivity of U.S. scientists (overall and across different types of institutions) to their policy environment, particularly policy-induced limits to their academic freedom to use particular inputs. Our evaluation of this changing policy environment in hESC addresses and clarifies a series of debates among policy-makers, including those who consider science policy in the light of national competitiveness, and the scientific community itself. More broadly, analysis of this policy environment illuminates some critical issues in the microeconomic and institutional foundations of knowledge accumulation (Mokyr 2002). Specifically, it provides insight regarding degree to which a

policy limiting scientific freedom shapes overall participation in a particular research line, has distributional consequences on the nature of those who particulate or transforms the distribution of collaboration (Aghion, Dewatripont and Stein 2008). We see these as critical issues because: While "Open Science" is widely recognized to play a fundamental role in the production of fundamental knowledge and follow-on production that builds on these foundations (Merton, 1973; Dasgupta and David, 1994; David 1998; Stephan, 1996), few formal analyses support our understanding of the impact of policies and practices on scientific progress. More comprehensive analysis accounts for the fact that the substance of science policy takes many forms: choices about the level of (and restrictions on) public funding, rules governing access to scientific research materials and data etc. Science policy may influence both the overall productivity and the direction of the scientific research enterprise. More subtly, policies may have important *distributional* consequences: For example, national-level policies may have an impact on national scientific advantage by changing the degree to researchers in close proximity to an original discovery are able to exploit a discovery more rapidly and more intensively than distant researchers. Within a particular national policy context, there may be additional distributional consequences that influence researchers in high status institutions compared to those in less well-funded institutions. For example, while reduced access to key inputs may lower accessibility and productivity for the *average* researcher, some researchers (e.g., those at leading institutions) may be able to circumvent such limits through their differential access to additional resources, the use of their social status to form collaborations etc.

We use a difference-in-difference approach to evaluate the impact of the human embryonic stem cell (hESC) policy environment in the period 1998 to 2008, focusing specifically on the impact of the shift from high levels of policy uncertainty 1998-2001 to the 2001 policy which provided clarification but with high levels of restrictions. Not simply an analysis of *levels* of hESC research, we explore the distributional impact of policies on hESC research focusing on several dimensions. First, we explore the direct impact of the policy on the spatial distribution of production – determining its impact on research in the U.S. relative to other countries (both in absolute and quality-weighted terms). Finding that U.S. science did fall behind foreign scientists but then rapidly caught up, we focus on two additional distributional dimensions that allow us to examine the organizational mechanisms

used by U.S. scientists to catch-up to the frontier: The relative response of those at high versus low status universities, and the impact on national versus global research collaboration.

Our approach pays particularly close attention to the issue of the counterfactual – as is well known in the policy evaluation literature, without a precisely defined counterfactual it is difficult to determine whether for example, the pattern of the U.S. falling behind other countries is simply a function of the globalization of science or caused by the U.S. policy environment. Likewise, if we see low status university researchers falling behind their colleagues within the U.S. this may again reflect predictable patterns of cumulative advantage. To resolve these issues we use several controls groups. First, we consider other forms of stem cell research (including the use of human adult stem cells, mouse embryonic stem cells and other animal-derived adult stem cells). However, there are some concerns that these areas may be important substitutes for hESC researchers and so we also include a set of control articles selected from the same journal issue as each of our hESC articles. (We call these the "Nearest Neighbor" articles (Furman and Stern, 2006), as they reflect the three articles immediately preceding and following our treatment articles in the journal in which each treatment article appears.) However, a key concern with this strategy is that the follow-on research trajectory of normal science is distinctive from "hot" areas of science (Kuhn, 1962). Thus, we select another high profile area of science – RNA interference[1] (RNAi) – that like hESC research was enabled by an important breakthrough in 1998 and proceeded to follow a trajectory of follow-on knowledge accumulation around the world. It is important to note that this area is not a substitute for hESC researchers, as work in this area required specific training that differs from that obtained by hESC researchers (and vice versa). This allows us to compare the patterns in hESC with another breakthrough rather to "normal science" which may be subject to quite different distributional dynamics (Kuhn, 1962).

To undertake this analysis we have assembled a dataset of 719 articles, which include 17 hESC root articles, 93 other stem cell root articles (OSC), 51 root RNAi articles, and 559 Nearest Neighbor root articles (approximately six for each hESC and OSC article). Our stem cell article sample is drawn from the list of articles identified by the June 2001 NIH

[1] Andrew Fire and Craig Mello were awarded the 2006 Nobel Prize in Physiology/Medicine for this research.

report, "Stem Cells: Scientific Progress and Future Research Directions," as seminal contributions; our RNAi sample is based on a list published by Ambion Inc., a company that manufactures and markets products related to RNAi-research. We consider the set of articles that cite the original articles in our treatment and control groups in the period through the end of 2007 as indicators of the nature and extent of follow-on research. (We refer to each of our original treatment and control articles as "root articles," and label the follow-on articles as "citing articles".) For each follow-on publication, we have information about the location and institutional affiliation of the researchers, as well as other publication and researcher characteristics. Our empirical examination focuses on two principal questions: First, we ask whether the geographic distribution of hESC research after the 2001 stem cell policy differs significantly from the pattern of regional agglomeration/dispersion realized by research into non-human and/or non-embryonic stem cell lines. Second, we ask investigate the impact of the Bush administration policy on the distribution and organization of human embryonic research in the United States, examining whether the policy shifted the locus of research to different institutions across the status hierarchy or whether it affected the extent and nature of collaboration.

Prior work on this subject has demonstrated a decline in the extent of hESC research published by US and non-US authors in the years following the Bush administration policy decision (Owen-Smith & McCormick, 2005). Our approach makes a number of contributions relative to this state of knowledge. First, our analysis makes clear that that the interpretation of these data depends greatly on the inclusion of matched controls and on the ability to take advantage of nuanced details about the institutional affiliations of authors. Human embryonic stem cell research and RNAi research are substantively different from "normal science" during the sample period in the sense that research built on these initial discoveries at a greater rate than on discoveries in similar journals at the same time.

With respect to the geographic distribution of scientific discovery, our counterfactual analysis suggests that, after a period of significant policy uncertainty, U.S.-based hESC research experienced a 35%-40% decline between 2001 and 2007, relative to such research outside the United States. However, the negative impact of the policy on U.S. research output appears to have been tightly concentrated in the years 2001 to 2004; the data show a recovery towards expected levels 2005 and 2007. This recovery period may be rooted in a

behavioral response by U.S. institutions and researchers. The negative impact of the policy on hESC output is substantially less among researchers at elite U.S. universities. Indeed, by 2004, the impact of the policy on hESC research has virtually disappeared among elite U.S. universities. Similarly, the negative impact of the policy on hESC output lower for papers that involve collaborations between U.S. authors and researchers outside the United States. In conjunction with qualitative findings, we interpret these results as implying (a) that researchers at elite U.S. institutions responded to the policy by developing alternate (non-federal) mechanisms to support hESC research and (b) that an alternative, and potentially complementary, method to continue contributing to hESC research, despite the funding restrictions, involved collaboration with researchers in less restrictive policy environments.

II. Institutional Details

II.1. The Scientific Breakthrough

In1998 James Thomson and his colleagues from the University of Wisconsin made an unexpected announcement with the publication of their manuscript in leading journal *Science* documenting the first ever isolation of human embryonic stem cells[2]. This advance built on progress in the 19th and early 20th century as progress in biology led to the observation that certain cells could produce other cells, most notably blood cells. The pace of stem cell research accelerated in the late 1960s with animal based research on in-vitro fertilization techniques 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. Thomson's research came 17 years after Evans and Kaufman, and Martin published two papers reporting the first isolation of mouse embryonic stem cells. Although the Wisconsin had been able to isolate monkey embryonic stem cells a two years earlier, their work with human cells was a surprise to many scientific observers and those in the scientific community. It was also hailed as a "Breakthrough of the Year" by *Science* in January 1999, and later regarded as a major stepping stone in the history of science. The University of Wisconsin described the work as "a new era of human biological

[2] 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.

research, providing scientists with "blank slate" cells capable of becoming any of the more than 200 specialized cells in the body and offering researchers a rare view into the earliest stages of human development" (Wisconsin http://www.news.wisc.edu/14806).

Unlike some other advances, human ESCs were immediately recognized as both a critical new research tool making a contribution to further developments in basic research and an understanding of embryology and to developments in applied research with the potential to lead to novel therapies – in other words they lie in Pasteur's Quadrant (Stokes 1997, Murray and Stern 2007). In the period that followed, biologists described stem cell research as one of the most promising areas for scientific progress with scientists around the world anticipating rapid advances using both embryonic and non-embryonic stem cells, as well as human and non-human stem cell sources.

The intense interest in stem cells is 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 conduct 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" (NIH, 2001, ES-1). Thus both the scientific and medical promise of stem cell research derives from their potential to develop multiple types of cells. It is important to note the distinction between adult stem cells and embryonic stem cells. "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. In the body, it too, can renew itself." (NIH, 2001, ES-1). Embryonic stem cells are pluripotent in other words; they can develop any of the over 200 cell types in the body. As a consequence, in the late 1990s and certainly until 2008 (when our analysis ends), leading scientists regarded embryonic stem cells as holding more promise as research tools and as medical therapies because their greater ability to differentiate.

II.2. U.S. Stem Cell Policy History Before 2001

Even in the decades prior to Thomson's development of hESCs there had been considerable debate over the appropriate levels of support for research on human embryos

with variation across national and local governments with, specifically, substantial variation within the U.S., over time. The U.S. federal government has often imposed restriction on its support for research involving human embryos and other aspects of conception (Fletcher, 2000). Beginning in 1973, US government policy 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. In the late 1980s, both the NIH and Congress became more sympathetic to the prospect of using government funding for research on human embryos. President Clinton initially supported the removal of the ban on federal funding, but reversed course early in his presidency. Congressional action in 1995 expressly prohibited the use of federal funding for the development of human embryos that would either be destroyed or employed in research. This limitation essentially precluded federal support for *in vitro* fertilization, which usually creates more embryos than are deployed. The policy environment became more supportive of the federal support for human embryonic stem cell research during the final years of the Clinton Administration however considerable uncertainty remained. 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 soliciting proposals for future research. Thus, the United States began 2001 amidst a contentious public debate and an evolving, uncertain policy environment in which the scientific community actively touted the prospective value of hESC research and remained hopeful but uncertain about the prospects for large-scale funding.

II.3. The 2001 Bush Administration Stem Cell policy

After the uncertainty that beset the Bush versus Gore election was resolved, the Bush Administration initiated an official review of its policy options with respect to human embryonic stem cell research. It also placed a hold on the funding of proposals solicited by the NIH. As part of the administration review process, in February 2001Tommy Thompson, the Secretary of Health and Human Services and former Governor of Wisconsin 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.

On August 2001 in an environment of substantial interest and speculation, President Bush introduced his administration's policy, a policy that was met with considerable surprise in the U.S. media and by scientists. The policy included three features that are notable for our current project: The policy (1) enabled federal funding for research on a set of human embryonic cell lines that were existing at the time of the policy, (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. Thus, researchers who wanted to conduct work on non-approved hESC lines and also receive federal support for research on approved lines were obliged to establish laboratories that are physically and organizational distinct from one another. On the other hand, the policy placed no federal restrictions on the funding of human adult stem cell research or animal research on either adult or embryonic stem cells. This nuanced policy thus provided the first large-scale federal funding for research on human embryonic stem cell research while also ensuring that the U.S. federal government would not financially support the destruction of human embryos. The policy also formally opened the opportunity for interested non-federal actors to support hESC research efforts. Overall, the U.S. federal government allocated approximately $550 million to stem cell research in 2005; only $24 million of this, however, was devoted to human embryonic stem cell research (Beardsley, 2005). As a caveat (not analyzed in our paper), 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 new human embryonic stem cell lines and research on these lines. 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.

II.4. The international context – cross-country comparisons

The policy enacted by the Bush Administration in 2001 was by no means the most restrictive of all national policies: At the time of the policy announcement, a number of European countries did not offer permission to scientists to derive stem cell lines, conduct research on existing lines, or research involving somatic cell nuclear transfer, including

Austria, Ireland, and Italy. The policy environment in Germany was also restrictive: National policies allowed research on existing lines, but prohibited nuclear transfer and the derivation of new lines. At the same time, some countries had a long history of supporting stem cell research to a substantially greater degree. Leading stem cell research countries included Israel, Singapore, Sweden, and the United Kingdom (as well as follower countries China, Japan, and South Korea). In most of these nations allowed research on existing cells and nuclear transfer and also supported the derivation of new stem cell lines. A summary of national policies with respect to stem cell research in the year 2005 appears in Appendix Table 1. Thus while the precise nuance of policies in each nation is complex, overall, the background to the U.S. policy changes is one of at least a sub-set of uniformly supportive nations in a position to move rapidly into human embryonic stem cell research.

III. Methods & Research Design: Identifying the Impact of a Nuanced Policy Shift On The Geography Of Scientific Advantage

III.1. An experimental approach to assessing the impact of science policy shocks

Our approach assessing the impact of the Bush Administration policy intervention exploits a number of features of the scientific system and the storage and ease of use of bibliometric data. It is premised on several assumptions. First, we believe that data on the production and citation of academic papers provide valuable (though imperfect) indicators of scientific progress. We rely on the seminal work of Merton (1973), Garfield (1955), and De Solla Price (1971) in articulating the importance of priority and publication in the system of scientific rewards and noting the importance of publications and citations in tracking the rate and direction of scientific progress.[3] We rely upon the fact that (a) academic papers are produced at a specific and measurable point in time and (b) the use by follow-on researchers of the knowledge articulated in those papers takes place over time and in a way that can be measured as well. We interpret citations to academic papers as evidence of the use of prior

[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 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. We take care 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.

knowledge by follow-on researchers, although we acknowledge that these are noisy measures.

Our second assumption is that the degree to which future research "draws upon" (cites) a given article (and by whom and where and when) depends on institutional mechanisms, including intellectual property rights over the knowledge disclosed in the article, rules and institutional arrangements governing access to research resources, and national and local policies. Further, the opportunity to take advantage of a given research trajectory by researchers in any one location or institution depends on access to funding, materials, and support infrastructure to conduct that research in a timely manner (i.e., before others are able to exploit the opportunity). The impact of institutions and policy interventions on facilitating this process of step-by-step scientific discovery is a key challenge for science policy, and a central focus of science policy analysis (Aghion Dewatrapoint & Stein 2005; Mokyr 2002).

Our third and final assumption is that science policy interventions that change the institutional environment for scientific research will be reflected in changes in the rate and direction of scientific progress, which in turn is captured in citation patterns.[4]

From an experimental perspective, the econometrician would ideally observe a given piece of knowledge in distinct institutional or policy environments and compare the impact of that knowledge across regimes. To do so, our analytical framework relies on the fact that institutional changes or policy interventions may induce changes in the production of scientific articles or changes in citation behavior relative to baseline levels.[5] Moreover this

[4] Given these assumptions, it should be possible to observe and evaluation the impact of science policies on scientific progress. In reality, social scientists and policy analysts face a considerable challenge in assessing the extent to which any particular institution or policy influences the way in which the "knowledge stock" is created, maintained, and extended. In particular, it is empirically difficult to separate the influence of a particular institution or policy from the influence of the knowledge in which it is embedded, even though the two are conceptually distinct. Specifically, a selection effect may result from a correlation between the characteristics of institutional and policy regimes and the type of knowledge associated with them giving rise to a fundamental inference problem. Specifically, for a given piece of knowledge produced or diffused within a given institutional or policy environment, one cannot directly observe the counterfactual impact that knowledge would have had if the knowledge had been produced and diffused in an alternative institutional or policy setting. Moreover, even it if were possible to evaluate the average impact of a particular policy, ideally we would like to know how such an intervention impacted particular sub-populations of scientists and its impact on both high and low quality research/researchers.

[5] There are, of course, some important caveats to this approach. First, not all research is disclosed in the scientific literature; indeed, for-profit entities may decline to publish research results either to increase the costs of rivals' research (Rosenberg, 1990) or in the event that such results are disadvantageous for the firm. Second, a increases citations (relative to a baseline) may occur not because of the increased importance of a particular

natural experiment approach exploits the fact that the institutional environment changes over time in ways that do not impact the original "piece of knowledge" but which do impact the incentives and opportunities for follow-on researchers to exploit that piece of knowledge in their own research.

III.2. Evaluating the impact of the Bush Administration hESC policy shock

Because the outcome of the November 2000 was particularly uncertain and the specifics of the Bush Administration were also uncertain prior to its announcement in August 2001, we interpret the policy as a (plausibly) exogenous shift in the policy environment. We then examine the impact of this shift by comparing the rate and nature of human embryonic stem cell research with that of other types of stem cell research, other important research in cell biology, and a more loosely-matched control sample of related science.

Specifically, we investigate two central issues whether there is evidence that the Bush Administration policy had an impact on the geography of stem cell research – i.e., whether the enactment of the policy is associated with a shift the relative level of US vs. Non-US human embryonic stem cell research – accounting for trends in the progress of related science. In addition, we examine the distributional consequences of the policy, investigating whether the policy led to a shift in the nature of hESC research in the United States relative to other locations. For example, we are interested in (a) the extent to which the policy affects researchers at the highest status institutions differently from researchers at other institutions and (b) the extent to which the policy affects collaboration among researchers.

To do this, we divide the citations from each citation year into mutually-exclusive types and estimate the impact of the HESC_POST_2001 on each citation-year margin. For example, to test whether the policy shift decreases the output of US HESC research relative to Non-US research, we estimate the difference of the impact of the policy shift on follow-on publications by Citing articles with US-based authors versus Citing articles with Non-US based authors. To do this, we estimate the joint equations:

'unit' of knowledge, but simply because of the ease of its availability relative to alternative pieces of knowledge or for other reasons (such as changes in author prominence or position) that do not reflect changes in the actual use of knowledge. Such problems would average out across the areas we study, unless these changes are closely correlated with the specific policy or institutional changes we study.

$$(1)\ US-CITES_{it} = f(\varepsilon_{it};\gamma_i + \beta_t^{US} + \delta_{t-pubyear}^{US} + \alpha^{US} + \psi_0^{US} 2001_{it} + \psi_0^{US} POST2001_{it})$$

and

$$(2)\ Non-US-CITES_{it} = f(\varepsilon_{it};\gamma_i + \beta_t^{Non-US} + \delta_{t-pubyear}^{Non-US} + \alpha^{Non-US})$$

where *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 2001 and POST-2001 represent dummy variables equal to one only in year 2001 and 2002-2007, respectively. The coefficients on 2001 (ψ_0) and POST-2001 (ψ_1) indicate the marginal impact of the intervention on the sets of treated articles. Since the only the articles in (1) are affected by the policy, we estimate the impact on the policy only on these sets of citations. This specification includes one key parametric restriction. We set the root article fixed effects (γ_i) to be equal across the two equations. We allow calendar year and article age fixed effects to vary across equations, however. By estimating these equations jointly (in "stacked regressions"), 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 on the relative to HESC in the rest of the world.

We can modify (1) and (2) to facilitate additional comparisons. For example, we could add additional equations ("stacks") to the joint estimations that will allow us to investigate the impact of the policy shift on additional citation margins.

In addition, by modifying (1) and (2) to allow for pre-deposit and post-deposit dynamics it is possible to estimate whether the impact of the policy changes with the time elapsed since policy intervention and to check for the presence of a pre-deposit time trend. The former is important to understand 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 levels or diffusion of knowledge, whether it declines or returns to baseline over time, or whether the policy intervention induces continuously growing effects. The pre-deposit trend might provide evidence about the exogeneity of the policy intervention itself.

While the prior paragraphs focus on the impact of policy interventions or institutional changes on the overall count of citations to a given discovery, we are also interested in how

the policy shift affects the relative distribution of hESC research by institution type, the quality of journals in which follow-on research appears, and the extent and nature of collaborative research. To estimate the impact of policy interventions on each of these subpopulations, we can aggregate these individual citations into counts of the number of citations received by a given article in a given year by a given subpopulation of citers.

To implement this approach we construct a dataset composed of scientific publications linked to four types of articles: (1) seminal human embryonic stem cell research articles, (2) seminal articles on all other types of stem cell research, (3) seminal RNAi articles, and (4) control articles matched to the sets described in (1) and (2). Because we observe citations to a scientific publication both before and after the policy shock (and because we are able to identify a counterfactual estimate of the citation rate that would have occurred had the shock *not* occurred), we can identify the causal impact of the Bush Administration policy shock on the pattern of citations to a scientific publication.

Of course, citations data takes the form of count data that are skewed to the right and over-dispersed relative to Poisson. As well, the rate of citation to a given piece of research will vary with the calendar year and with the time elapsed since initial publication. Therefore, except where noted, we employ a conditional negative binomial model with age and year fixed effects for citations produced per year for each scientific article in our dataset.[6,7] We experiment with a range of alternative specifications.

IV. Data

To implement the difference-in-differences framework we articulate above, we identify multiple series of treatment and control articles, and construct a database that

[6] Several subtle 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.

[7] When using a conditional fixed effects estimator, one citation year and one age fixed effect are not separately identified (Hall et al, 2005). Since the main effect that we are interested in is separable from these effects, the precise specification we employ to overcome this identification issue does not at all affect our estimate of the impact of BRC deposit on citations. In our estimation, we identify differences relative to age = 0, and relative to publication in years after 1975 (though, due to data limitations, we actually impose a single regressor on the years 1975-1979).

includes bibliometric information on both the original articles and ("root articles") the follow-on articles that cite them ("citing articles").

Our sample of stem cell articles is comprised of the publications identified by the NIH report, "Stem Cells: Scientific Progress and Future Research Directions." This report was published in June 2001 and was an input into the Bush Administration policy-making process. The report was devoted to scientific facts relevant to the policy debate, but does not appear to be a political document. 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). Most importantly, the document identifies 110 articles that reflect the seminal articles in stem cell research, including paper associated with embryonic and adult stem cells derived from both human and animal models. We consider these as our root articles, and track patterns of forward citations received by these articles. Of these articles, 17 root articles are associated with pioneering work in human embryonic stem cell research. We consider these to be our primary treatment sample. We consider the remaining 93 NIH publications, which consist of embryonic and adult animal and human adult stem cell research articles, to be control articles. We recognize, however, that the U.S. policy shock may also affect the incentives and ability to contribute to work in these other areas of stem cell research.

Our second control sample consists of a set of seminal articles on RNA interference (RNAi), another area of cell biology, which, like human embryonic stem cell research, experienced a substantial breakthrough in 1998 pioneered by U.S.-based researchers. As our sample of RNAi root articles, we employ a 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.[8] We consider this sample to be a particularly valuable control sample, as it represents a scientific breakthrough that (a) was achieved in essentially the same field and at the same time as hESC research, (b) was of similar (or, indeed, greater

[8] Our results are not sensitive to including or omitting the 4 RNAi articles published in 2002.

scientific importance – Andrew Fire & Craig Mello won the 2006 Nobel Prize for their 1998 work in RNAi), and (c) was also introduced in the United States.

Our third set of control articles are less well-matched with respect to scientific importance but are more precisely matched with respect to publication timing and journal. Specifically, our third control sample consists of "nearest neighbor articles," including each of the three articles that immediately precede and follow each of the root stem cell articles in the same year and issue of the journal in which the stem cell article was published (Furman and Stern, 2006). Thus, each stem cell article in our sample includes up to six nearest neighbor control articles.

V. Results

V.A. Sample Description

Our data includes 719 root articles, of which 17 are human embryonic stem cell root articles, 93 are other stem cell root articles, 56 are RNAi root articles, and the remainder are nearest neighbor articles. We report descriptive statistics for the root articles in Table 1; Panel A presents data for the entire sample, while Panel B decomposes the data by article type. Across the sample, more than 60 percent of root articles include a reprint author based in the United States; approximately 50% of hESC reprint authors are US-based.[9] More than two-thirds of Reprint Authors' addresses identify university affiliations, while 8% identify non-university hospital affiliations, and only 2% can be linked to firms. These figures sum to less than one, because some Reprint Author data do not allow assignment into institutional categories. We classify 24 percent of reprint authors as associated with "Top 25" universities, based on the Center for Measuring University Performance (Arizona State University) 2006 Annual Report of university research rankings and 20.2% to be associated with "Top 50" universities. Based on a definition that considers "top journals" to be journals that achieve an ISI Journal Citation Impact greater than 25, we find that citations to root articles published in top journals constitute 75% of the sample.

Sample articles receive 17.3 citations per year, with a standard deviation of 39.4. hESC articles receive approximately 32 citations per year. Consistent with the belief that

[9] Another convention that we could use to determine the geographic location of authors is one based on the complete set of institutions and addresses listed in the ISI field "C1". We have begun to experiment with this field. Using this field enables a single paper to be associated with multiple countries-of-origin.

RNAi constitutes a significant advance in biology, the RNAi root sample receives even more citations than either of the stem cell samples or the nearest neighbor articles (mean = 68.0). Relative to either category of stem cell article, RNAi articles include more citations from US-based authors. The Nearest Neighbor sample is the least well-cited, receiving, on average, fewer than 9 citations annually.

V.B. 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 noticably 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 US and non-US 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 US-based author or (b) no US-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, 2006). The rate of obsolescence is of less interest in our analysis than the relative levels of US-based and non-US-based citations.

The top-most graphic in Figure 2 compares trends in citations to hESC root articles by US-based and non-US-based Reprint Authors. While the number of citations by each category is similar between 1998 and 2000, the counts diverge beginning in 2001. Specifically, while the growth rate of non-US-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 (2005), who conclude based on a keyword approach to identifying hESC publications up until l2004 that the US share of hESC articles experienced a relative decline beginning in 2001. The US 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. US-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 US share of RNAi and Nearest Neighbor articles, the unambiguous relative decline of the US share of hESC articles and modest decline in other stem cell articles may suggest that the shock to US funding policy had an impact on the rate of follow-on stem cell research in the United States. The results appear to be far from unambiguous, however, as the relative decline in US hESC share begins in 2001, although the policy was not introduced until August of that year. These 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 the regression analyses that follow, we attend to each of these issues.

V.C. Core Results

Our empirical approach relies on a differences-in-differences analysis using matched control samples that attempts to isolate the impact of the U.S. policy shock on the rate and nature of follow-on research. This strategy requires observing research articles in two distinct policy environments, associated with a pre-shock and post-shock period. A pre-requisite for assessing the impact of the policy shock on hESC research is identifying the most thoughtfully-matched treatment and control samples on which to base our counterfactual analyses. In our principal analyses, we rely on RNAI research as our control group. Thus, we derive our counterfactual estimate of the impact of the Bush Administration policy shift 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. Although some of our extended analyses also compare the geographic evolution in hESC research to that of Other Stem Cells and Nearest Neighbors. 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, however, as the techniques and materials vary greatly. The nearest neighbor articles also form a less satisfactory control group, as publications, even within a journal, may vary in content and kind across fields and, thus, may have substantially different citation profiles.

We present our regression results beginning on Table 3 with a conditional fixed effects negative binomial specification in which we estimate $Citations_{it}$ as the dependent variable. (OLS would be inappropriate for inference in this context, as our citation data are composed of highly skewed count data.) To implement our approach, we stack the entire set of observations on top of itself. In the top stack, the dependent variable takes the values of $Non\text{-}US\text{-}Citations_{ij}$ and the dependent variable takes the values of $US\text{-}Citations_{ij}$ in the bottom stack. Each stack includes dummy variables for HESC_2001 (hESC publications in year 2001) and HESC_POST_2001 (hESC publications in years 2002-2007). The second stack, in which the dependent variable reflects the number of $US\text{-}Citations_{ij}$ also includes HESC_2001_US and HESC_POST_2001_US, which enable us to estimate the additional boost (or decrement) to citations that HESC root articles receive in 2001 and 2002-2007 in the US relative to the rest of the world. In all specifications, we include (but do not report) a set of common article and article age fixed effects and separate calendar year fixed effects for each stack, which enable trends in citation patterns to vary between the US and rest of the world. All models should include block bootstrapped standard errors, clustered by article [MacKinnon, 2002]; however, the current set of regressions does not correct the standard errors in this way and, thus, overstates the statistical significance of each estimated coefficient. 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. We 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).

Columns (3-1) and (3-2) estimate the impact of the policy shift on citations in year 2001, the year of the shock, and in the average of the years following the shock (2001-2007). The columns differ in that the dependent variable in (3-1) considers citations to be US-based if any address in the address field is US-based, while (3-2) considers citations to be US-based only if the Reprint Author is US-based. In both equations, the coefficients on the variables HESC_2001 and HESC_POST_2001 describe the average difference in citations between HESC and RNAI articles in the years 2001 and 2001-2007, respectively, controlling for year, article age, and article-specific fixed effects. The magnitude and lack of statistical significance of these coefficients in both columns suggest that HESC and RNAI research grew at relatively similar rates during the post-shock period. The coefficients on HESC_2001_US and HESC_POST_2001_US indicate the increment (or decrement) to citations with any US address during the years 2001 and 2001-2007, respectively. 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 US address declined following the shock. The impact of the shock appears to are greatest in 2001, during which research output falls by more than 50%. In the years after the shock, the production of papers with any US-based author declines by 37%, while the production of papers with a US-based Reprint Author declines by 41%.

In order to better understand the dynamic impact of the shock, we decompose the impact of the shock by year of post-shock impact in (3-3) and (3-4). In these models, the coefficients on HESC_YEAR indicate the relative growth of HESC in comparison to RNAI during that year, while the coefficients on HESC_YEAR_US compare the output of HESC publications in that year in the US to HESC publications outside the US. These results suggest that the impact of the shock was most negative and severe in 2001, but that a recovery begins in 2004. By the end of the sample period, the negative impact of the policy on US-based HESC output has declined to less than 25 percent. The statistical significance of the difference has also declined, although some of this may be due to obsolescence and truncation.

While the econometric results are clear and are robust to a number of alternative specifications, the interpretation of these results is complex: There is an unambiguous decline in US-based HESC output relative to the rest-of-the-world, which begins in 2001 and continues to the end of our sample period in 2007. As the policy was not unveiled until August 2001 and the most severe decline in US HESC output occurred in the same year, the data do not support the interpretation that the enactment of the policy itself caused the decline. The data are, however, consistent with an interpretation in which changes beginning in 2001 (which could include scientist expectations regarding the policy shock, as well as a host of other factors), led to a relative decline in US HESC output. The incremental though not complete recovery in US HESC output over the sample period, particularly between 2004 and 2007, suggests a capacity for adaptation within the US HESC research community. There are a number of possibilities to explain this recovery. We investigate variation in the impact of the policy shock and a few potential explanations for the recovery in the analyses that follow.

In Table 4 we explore the robustness of the core results to variations in the nature of citations and the sample. Column (4-1) replicates the analysis of (3-1), using counts of citations in Top Journals rather than all journals as the dependent variable. The results imply that the policy shift had an even greater effect on US-hESC publications in top journals (47 percent decline) than that across all scientific journals (37 percent decline). The remaining columns in Table 4 address the possibility that the results in Table 3 are predicated on specific features of the RNAI control sample. In (4-2), we replicate the analysis of (3-1), excluding RNAI articles in the highest decile of citations received prior to the policy shift.[10] The principal results are unaffected by censoring the RNAI sample in this way. In (4-3), we substitute the nearest neighbor article sample for the RNAI sample as the control group. The positive and significant coefficients on HESC_2001 and HESC_POST_2001 indicate that, relative to the nearest neighbor controls ("normal science"), hESC ("hot science") experiences a nearly 60 percent and 120 percent boost in citations in 2001 and 2002-2007, respectively. The coefficients on HESC_2001_US and HESC_POST_2001_US, imply that

[10] Similar results obtain when we omit the top two declines of RNAI article in citations received prior to the policy shift. An improved way of ensuring that the RNAI sample serves as a suitable control for the hESC root articles would be to match RNAI and hESC root articles (to the greatest extent possible) on publication year, journal, and citations received before the policy shift. To this point, we have not had the opportunity to implement a matching procedure, although we intend to do so.

the falling-behind effect is also evident when comparing the hESC sample to the nearest neighbor root articles. Although the average treatment effect is statistically and economically significant in (4-1), (4-2), (4-3), the temporal pattern in (3-3) and (3-4) is also evident: decomposing the HESC_POST_2001 effect for the regressions in the first three columns of Table 4 yields the same results as in the prior table. In each case, the negative impact of the policy shift is most evident in 2001-2004, and there is evidence of a relative rebound in the final few years of the data. Each of these cases is 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 (4-4), 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 (although one infer some relative decline in HESC output in the year of the policy shift).

V.D. Extended Analyses – Mechanisms of Response

In Tables 5 and 6, we examine two potential mechanisms that may partially explain the relative rebound in US hESC output in the latter years of our sample. Specifically, Table 5 investigates the relative impact of the policy shift on elite US universities relative to other institutions, while Table 6 investigates the impact of the policy shift on various types of research collaborations. Table 5 compares the impact of the policy shock across three types of institutions: (a) elite US universities, (b) all other US institutions, and (c) institutions outside the United States. 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 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. Columns (5-1) and (5-2) examine the impact of the policy shift on citations in all ISI-indexed journals, while (5-3) and (5-4) track citations in top journals only. In these models, we have "triple-stacked" the data; coefficients on HESC*US*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

US in to the policy shift, each relative to citations from articles with Non-US citing authors. The average treatment effects imply a decline of approximately 33 percent of output in Top-25 US universities and 46 percent among non-elite university. The year-by-year effects, however, imply that HESC output by elite US universities declined in 2002 and 2003, but recovers nearly completely thereafter. By contract, HESC research output by US-based reprint authors in other institutions declines 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.

We examine an additional mechanism by which US-based HESC researchers may have responded to the policy shock in Table 6. Specifically, Table 6 compares follow-on research across three collaboration-location types: (a) papers with only US authors, (b) papers with US and non-US authors (i.e., those with international collaboration), and (c) papers with no US authors. The results suggest that the output of papers with only US-based authors declined more significantly than those of the other types following the policy shock. The output of HESC papers involving collaboration between US and non-US authors declined significantly in the year of the policy shock (2001), but are unaffected by the policy shock thereafter. These results suggest a behavioral response by scientists to the policy shock, similar in spirit to that observed by Murray (2009). These are limited results, but they are consistent with an explanation in which researchers (possibly those at the more resource-constrained institutions) in the United States collaborate with scientists outside the United States who may have access to resources more difficult to obtain in the United States.

VI. Discussion

In this paper, we present preliminary analyses of the impact of the Bush Administration Policy regarding Human Embryonic Stem Cell research on the competitiveness of US Stem Cell science in relation to that of the rest of the world. Our results suggest a modest relative decline in US-based work that builds on seminal hESC

research. The timing of the shift, however, suggests that the decline in US hESC output corresponds to the period in which uncertainty about the direction of the policy was resolved – i.e., when the result of the 2000 U.S. presidential election was known, even if may have begun before the 2001 policy was announced. This decline in US hESC research leadership is based on the comparison to the RNAI control sample; however, it arises even when using alternative ways of measuring scientific output and in using alternative the control samples. These results are consistent with evidence documenting the broad-based globalization of scientific and technical capabilities (Furman & Hayes, 2006; Hayes, 2008).

Interestingly, however, that the relative decline in US hESC output abates, beginning in 2004. Considering the policy restrictions placed on hESC research in the United States and the policy support provided in a number of other countries, including South Korea, Singapore, and Israel, the bigger surprise in our results may be the robustness of the United States hESC research community to modest perturbations of the system. US researchers at elite institutions were able to overcome federal funding limitations, principally, as our qualitative research suggests (and preliminary evidence from funding data shows), as a result of their ability to fund their research efforts with private sources of funding. The volume of research generated by researchers at institutions outside the elite circle did decline; however, there is evidence of valuable adaptation by US researchers. In particular, US-based researchers appear able to overcome funding difficulties by collaborating with researchers outside the United States. Although our results evidence a modest relative decline in U.S. hESC output, we thus interpret the overall findings as consistent with a picture in which the regional institutions that support scientific competitiveness are robust and relatively enduring. Overall, the impact of the Bush Administration funding restrictions may be of second-order importance relative to issues such as the extent of overall funding for the NIH (Stephan, 2008).

REFERENCES

Aghion, P., M. Dewatripont, and J.C. Stein (2005) "Academic Freedom, Private-Sector Focus, and the Process of Innovation."

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

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(2): 15-21.

De Solla Price (1971) *Big Science, Little Science.* 4th Edition.

Fletcher J.C. (2000) "Deliberating incrementally on human pluripotential stem cell research," *Ethical Issues in Human Stem Cell Research*, Vol. II. Commissioned Papers. National Bioethics Advisory Commission, US Government Printing Office: Rockville, MD, E1-E50.

Furman, J.L. and S. Stern (2006) "Climbing Atop the Shoulders of Giants: The Impact of Institutions on Cumulative Research," NBER WP# 12523.

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

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

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.

Merton, R. (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. (2007) "The Stem-Cell Market – Patents andd the Pursuit of Scientific Progress," *The New England Journal of Medicine*, 356(23): 2341-2343.

Murray, F. (2009) "The Oncomouse that Roared: Hybrid Exchange Strategies as a Source of Productive Tension at The Boundary Of Overlapping Institutions," *American Journal of Sociology.*

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

Owen-Smith, J. and J. McCormick (2006) "An International Gap in Human Embryonic Stem Cell Research," *Nature Biotechnology*, 24(4):391-392.

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

Schubert A. and T. Braun (1993) "Reference standards for citation based assessments," *Scientometrics*, 26(1): 21-35.

Thomson, J.A. Itskovitz-Eldor, J., Shapiro, SS, et al. (1998) “Embryonic stem cell lines derived from human blastocysts,” *Science* 282, 18271145-1147

Table #1 – Descriptive Statistics – Root Articles

Panel A: Characteristics of Root 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 based in Hospital	0.06	0.24	0	1
Reprint Author based in Firm	0.02	0.12	0	1
Reprint Author at Top 25 US University	0.24	0.43	0	1
Paper in Top Journal	0.75	0.43	0	1
Collaboration - any type	0.94	0.24	0	1
Collaboration - US-EU	0.10	0.29	0	1
Collaboration - US-Asia	0.03	0.16	0	1
Collaboration - university-firm	0.04	0.19	0	1

***** *Note that Reprint Author data do not allow assignment of all articles into institutional categories.*

Panel B: Characteristics of Root Articles, by article type

	Other Stem Cells *(n=93)*		**HESC** *(n=17)*		**RNAI** *(n=51)*		**Nearest Neighbors** *(n=559)*	
	Mean	**Std. Dev.**	**Mean**	**Std. Dev.**	**Mean**	**Std. Dev.**	**Mean**	**Std. Dev.**
Publication Year	1997.11	4.41	1995.19	6.01	1999.84	2.61	1997.36	3.81
Reprint Author based in US	0.56	0.50	0.53	0.49	0.76	0.43	0.61	0.49
Reprint Author based in Uni	0.69	0.47	0.87	0.35	0.56	0.50	0.68	0.47
Reprint Author based in Hosp	0.06	0.23	0.00	0.00	0.04	0.21	0.07	0.25
Reprint Author based in Firm	0.04	0.20	0.00	0.00	0.00	0.00	0.01	0.12
Reprint Author at Top 25 US Uni	0.19	0.39	0.27	0.46	0.20	0.40	0.25	0.43
Paper in Top Journal	0.73	0.45	0.63	0.50	0.96	0.20	0.74	0.44

Table #2 – Descriptive Statistics – Characteristics of Annual Citations Received

Panel A: Characteristics of Annual Citations Received, entire sample

Variable	Obs	Mean	Std. Dev.	Min	Max
Citing Year	6847	2002.54	3.06	1996	2007
Annual Citations	6845	17.26	39.36	0	698
Annual citations received from articles with…					
Reprint Author based in US	6845	7.33	16.66	0	305
Reprint Author not based in US	6845	9.93	23.65	0	516
Any author in US	6845	8.30	18.46	0	344
No author in US	6845	8.95	21.74	0	477
Reprint Author based in University	6845	11.12	25.65	0	487
Reprint Author based in Hospital	6845	0.89	2.48	0	32
Reprint Author based in Firm	6845	0.30	1.15	0	29
Reprint Author at Top 25 US University	6845	2.39	5.55	0	95
Paper in Top Journal	6845	6.22	15.35	0	360
Collaboration - any type	6845	15.51	35.98	0	672
Collaboration - US-EU	6845	1.25	2.78	0	51
Collaboration - US-Asia	6845	0.52	1.55	0	27
Collaboration - university-firm	6845	0.48	1.42	0	28

Panel B: Characteristics of Annual Citations Recieved, by article type

	Other Stem Cells *(n=901)*		***HESC*** *(n=159)*		***RNAi*** *(n=403)*		***Nearest Neighbor*** ***(n=5382)***	
Variable	**Mean**	**Std. Dev.**	**Mean**	**Std. Dev.**	**Mean**	**Std. Dev.**	**Mean**	**Std. Dev.**
Citing Year	2002.47	3.09	2002.32	3.19	2003.33	2.68	2002.50	3.07
Annual Citations	43.18	58.45	32.03	55.05	67.95	94.61	8.69	16.01
Annual citations received from articles with…								
Reprint Author based in US	16.70	22.50	11.06	19.43	31.53	41.80	3.84	7.44
Reprint Author not based in US	26.48	37.66	20.97	36.19	36.42	54.31	4.85	9.35
Any author in US	19.13	24.99	12.83	22.02	34.49	46.24	4.40	8.26
No author in US	24.05	34.95	19.19	33.44	33.47	49.67	4.29	8.48
Reprint Author based in Uni.	28.09	39.96	20.26	35.15	41.39	59.42	5.74	11.01
Reprint Author based in Hosp.	3.05	4.83	1.90	3.54	1.67	3.34	0.44	1.28
Reprint Author based in Firm	0.58	1.22	0.75	1.58	2.02	3.49	0.11	0.42
Reprint Author at Top 25 US Uni.	5.25	7.22	4.14	7.66	10.18	13.64	1.27	2.75
Paper in Top Journal	11.41	14.08	6.70	10.74	32.65	46.06	3.35	6.31
Collaboration - any type	38.69	53.93	26.99	46.88	61.20	87.64	7.87	14.61
Collaboration - US-EU	2.76	3.74	2.01	3.41	4.16	6.62	0.76	1.61
Collaboration - US-Asia	1.28	2.29	1.03	2.32	2.11	3.93	0.26	0.73
Collaboration - university-firm	1.32	2.19	1.26	2.49	1.80	3.26	0.22	0.67

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

Stacked Negative Binomial Regressions; IRRs reported; SEs not bootstrapped (& not adjusted to IRRs)

Sample includes hESC and RNAI roots articles only	**(3-1)**	**(3-2)**	**(3-3)**	**(3-4)**
	DV = Cites with Any US address (or No US Address)	**DV = Cites with US Reprint Author (or Not US Reprint Author)**	**DV = Cites with Any US address (or No US Address)**	**DV = Cites with US Reprint Author (or Not US Reprint Author)**
HESC_2001	1.132	1.081		
	(0.203)	(0.190)		
HESC_POST_2001	1.104	1.079		
	(0.129)	(0.125)		
HESC_2001_US	**0.491**	**0.488**		
	(0.117)* **	**(0.118)*** **		
HESC_POST_2001_US	**0.631**	**0.589**		
	(0.046)* **	**(0.045)*** **		
HESC_1996_US			1.522	1.197
			(0.719)	(0.706)
HESC_1997_US			1.140	1.057
			(0.498)	(0.646)
HESC_1998_US			0.801	0.855
			(0.335)	(0.361)
HESC_1999_US			0.864	0.635
			(0.294)	(0.221)
HESC_2000_US			1.021	0.891
			(0.274)	(0.247)
HESC_2001_US			**0.493**	**0.491**
			(0.108)* **	**(0.113)*** **
HESC_2002_US			**0.521**	**0.425**
			(0.095)* **	**(0.084)*** **
HESC_2003_US			**0.434**	**0.402**
			(0.072)* **	**(0.072)*** **
HESC_2004_US			**0.686**	**0.667**
			(0.098)* **	**(0.102)*** **
HESC_2005_US			**0.761**	**0.737**
			(0.115)*	**(0.119)***
HESC_2006_US			**0.704**	**0.659**
			(0.110) **	**(0.111)** **
HESC_2007_US			**0.749**	0.793
			(0.123)*	(0.140)
HESC_1996			1.380	1.332
			(0.579)	(0.469)
HESC_1997			1.843	1.534
			(0.699)	(0.512)
HESC_1998			**1.914**	1.507
			(0.682)*	(0.520)
HESC_1999			**1.764**	1.591
			(0.598)*	(0.507)
HESC_2000			1.072	0.905
			(0.321)	(0.259)
HESC_2001			**1.730**	1.343
			(0.471) **	(0.348)
HESC_2002			1.381	1.144
			(0.368)	(0.287)
HESC_2003			**1.625**	1.275
			(0.439)*	(0.327)
HESC_2004			**1.655**	1.287
			(0.435)*	(0.323)
HESC_2005			**1.902**	1.480
			(0.523) **	(0.389)
HESC_2006			**2.106**	**1.653**
			(0.582)* **	**(0.432)***
HESC_2007			**1.982**	1.474
			(0.564) **	(0.406)
Observations (articles)	1124 (67)	1124 (67)	1124 (67)	1124 (67)
Log Likelihood	-3202.34	-3187.15	-3177.63	-3161.86

Models include unreported Stack-specific Year FEs, Stack FEs, Article Age FEs, and Article FEs.
*Standard errors in parentheses / * significant at 10%; ** significant at 5%; *** significant at 1%*

Table 4: Additional Results – – US hESC Output vs. Rest-of-World (1996-2007)

Stacked Negative Binomial Regressions; IRRs reported; SEs not bootstrapped (& not adjusted to IRRs)

	(4-1)	**(4-2)**	**(4-3)**	**(4-4)**
	Sample: hESC & RNAI root articles	**Sample:** RNAI articles, censored to exclude Top 10% most highly-cited articles	**Sample:** hESC & Nearest Neighbor root articles	**Sample:** hESC & Other Stem Cell root articles
	DV = Cites in Top Journals with Any US address (or No US Address)	**DV = Cites with Any US address (or No US Address)**		
HESC_2001	1.142	1.124	**1.590**	1.041
	(0.258)	(0.208)	**(0.234)*** **	(0.161)
HESC_POST_2001	1.185	1.137	**2.190**	0.875
	(0.181)	(0.136)	**(0.206)*** **	(0.084)
HESC_2001_US	**0.506**	**0.500**	**0.559**	**0.656**
	(0.145)**	**(0.122)*** **	**(0.115)*** **	**(0.141)****
HESC_POST_2001_US	**0.532**	**0.624**	**0.669**	0.893
	(0.056)* **	**(0.048)*** **	**(0.046)*** **	(0.064)
Observations	1124	1050	10964	2102
Number of rart_num	67	62	569	108
Log Likelihood	-2425.25	-2869.98	-18013.49	-5686.58

*Models include HEAC*YearFEs, Stack-specific Year FEs, Stack FEs, Article Age FEs, and Article FEs.*
*Standard errors in parentheses / * significant at 10%; ** significant at 5%; *** significant at 1%*

Table 5: Mechanisms of Recovery – hESC research output by Elite and Non-Elite US universities

Triple Stacked Negative Binomial Regressions; IRRs reported; SEs not bootstrapped (& not adjusted to IRRs)

	(5-1)	(5-2)	(5-3)	(5-4)
	DV = Cites from US-Top-25 Uni-Authors, US-Not-Top-25-Uni, and Non-US-RP-authors		DV = Cites in Top Journals from US-Top-25 Uni-Authors, US-Not-Top-25-Uni, and Non-US-RP-authors	
HESC 2001	1.127 (0.180)		1.181 (0.247)	
HESC post-2001	1.141 (0.119)		1.245 (0.177)	
HESC 2001 (Top 25 University RP author)	0.703 (0.210)		**0.984** **(0.330)**	
HESC post-2001 (Top 25 University RP author)	**0.670** **(0.062)*****		**0.656** **(0.084)*****	
HESC 2001 (US Not Top 25 University RP author)	**0.412** **(0.103)*****		**0.373** **(0.118)*****	
HESC post-2001 (US Not Top 25 University RP author)	**0.537** **(0.041)*****		**0.465** **(0.054)*****	
HESC_1996 & Top-25 US University		0.356 (0.407)		0.000 (0.001)
HESC_1997 & Top-25 US University		1.643 (1.287)		3.849 (5.984)
HESC_1998 & Top-25 US University		0.946 (0.562)		0.578 (0.682)
HESC_1999 & Top-25 US University		0.786 (0.324)		**0.190** **(0.122)*****
HESC_2000 & Top-25 US University		0.876 (0.293)		0.934 (0.350)
HESC_2001 & Top-25 US University		0.699 (0.198)		1.021 (0.311)
HESC_2002 & Top-25 US University		**0.408** **(0.103)*****		0.713 (0.190)
HESC_2003 & Top-25 US University		**0.409** **(0.091)*****		0.656 (0.185)
HESC_2004 & Top-25 US University		0.871 (0.162)		0.830 (0.216)
HESC_2005 & Top-25 US University		0.826 (0.165)		0.584 (0.171)*
HESC_2006 & Top-25 US University		0.835 (0.165)		0.547 (0.197)*
HESC_2007 & Top-25 US University		0.993 (0.214)		0.684 (0.204)
HESC_1996 & Not Top 25 US University		1.796 (1.132)		0.000 (0.001)
HESC_1997 & Not Top 25 US University		0.801 (0.579)		3.849 (5.057)
HESC_1998 & Not Top 25 US University		0.788 (0.349)		1.050 (0.710)
HESC_1999 & Not Top 25 US University		**0.540** **(0.197)***		**0.462** **(0.175)****
HESC_2000 & Not Top 25 US University		0.909 (0.253)		1.116 (0.347)
HESC_2001 & Not Top 25 US University		**0.415** **(0.097)*****		**0.388** **(0.111)*****
HESC_2002 & Not Top 25 US University		**0.429** **(0.085)*****		0.647 (0.149)*
HESC_2003 & Not Top 25 US University		**0.388** **(0.070)*****		**0.382** **(0.107)*****
HESC_2004 & Not Top 25 US University		**0.575** **(0.088)*****		**0.355** **(0.083)*****
HESC_2005 & Not Top 25 US University		**0.656** **(0.105)*****		**0.520** **(0.126)*****
HESC_2006 & Not Top 25 US University		**0.570** **(0.096)*****		0.736 (0.199)
HESC_2007 & Not Top 25 US University		**0.648** **(0.114)****		**0.321** **(0.104)*****
Observations (articles)	1686 (67)	1686 (67)	1686 (67)	1686 (67)
Log Likelihood	-4268.57	-4232.29	-3248.90	-3222.80

Table 6: Mechanisms of Recovery: Collaboration between US & Non-US authors

Triple Stacked Negative Binomial Regressions; IRRs reported; SEs not bootstrapped (& not adjusted to IRRs)

	(6a-1)	**(6a-2)**
	DV = Cites by (1) Papers with Only US addresses; (2) US address & Non-US addresses ; & (3) Cites by Papers with No US addresses	
HESC_2001	1.174	
	(0.189)	
HESC_POST_2001	1.164	
	(0.122)	
HESC_2001_US-Only	**0.548**	
	(0.119)* **	
HESC_POST_2001_US-Only	**0.533**	
	(0.037)* **	
HESC_2001_US-&-Non-US Collab	**0.216**	
	(0.101)* **	
HESC_POST_2001_US-&-Non-US Collab	**1.001**	
	(0.096)	
HESC_1996_US-Only		1.060
		(0.515)
HESC_1997_US-Only		0.879
		(0.389)
HESC_1998_US-Only		0.736
		(0.307)
HESC_1999_US-Only		0.777
		(0.249)
HESC_2000_US-Only		1.000
		(0.250)
HESC_2001_US-Only		**0.555**
		(0.111)* **
HESC_2002_US-Only		**0.416**
		(0.074)* **
HESC_2003_US-Only		**0.352**
		(0.057)* **
HESC_2004_US-Only		**0.635**
		(0.086)* **
HESC_2005_US-Only		**0.669**
		(0.097)* **
HESC_2006_US-Only		**0.612**
		(0.092)* **
HESC_2007_US-Only		**0.637**
		(0.101)* **
HESC_1996_US & Non-US Collaboration		3.893
		(2.554)**
HESC_1997_US & Non-US Collaboration		3.223
		(2.307)
HESC_1998_US & Non-US Collaboration		0.931
		(0.604)
HESC_1999_US & Non-US Collaboration		1.110
		(0.543)
HESC_2000_US & Non-US Collaboration		0.965
		(0.404)
HESC_2001_US & Non-US Collaboration		**0.197**
		(0.090)* **
HESC_2002_US & Non-US Collaboration		1.060
		(0.268)
HESC_2003_US & Non-US Collaboration		0.829
		(0.191)
HESC_2004_US & Non-US Collaboration		0.885
		(0.184)
HESC_2005_US & Non-US Collaboration		0.989
		(0.199)
HESC_2006_US & Non-US Collaboration		0.969
		(0.201)
HESC_2007_US & Non-US Collaboration		1.029
		(0.222)
Observations (articles)	1686 (67)	1686 (67)
Log Likelihood	-4216.35	-4180.43

Figure #1 – Citing articles by year

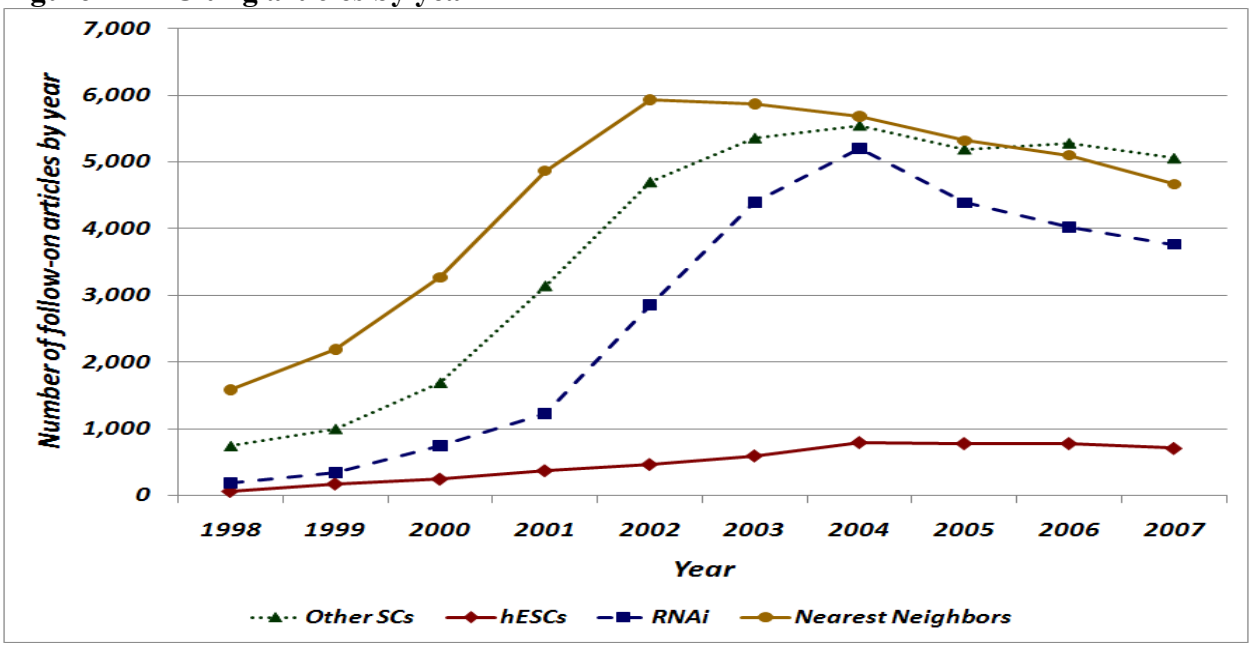
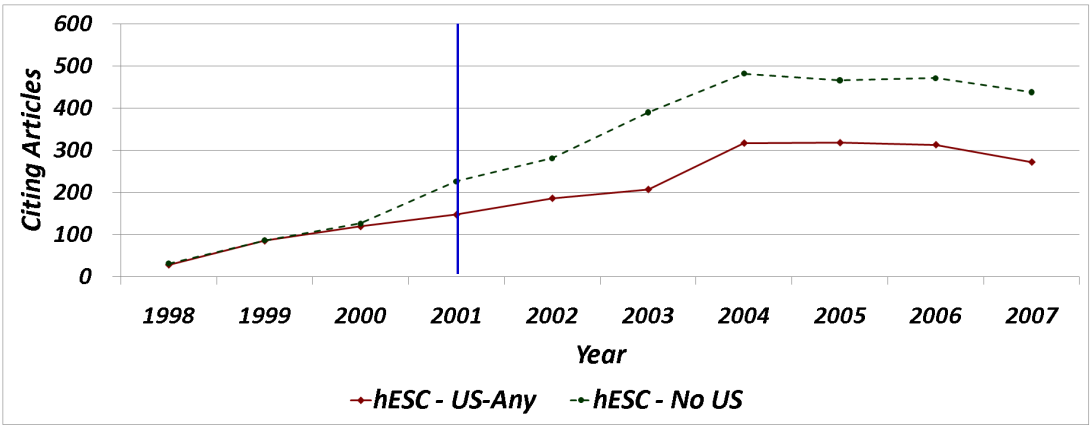
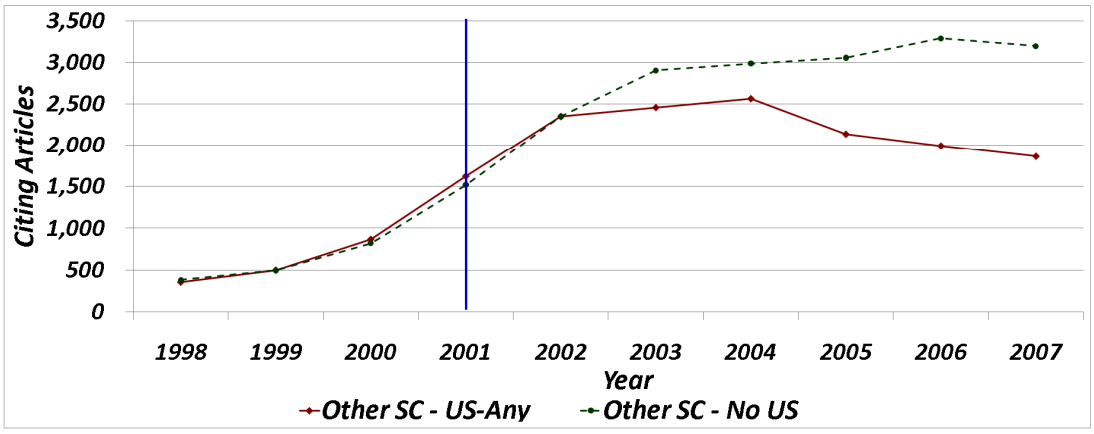
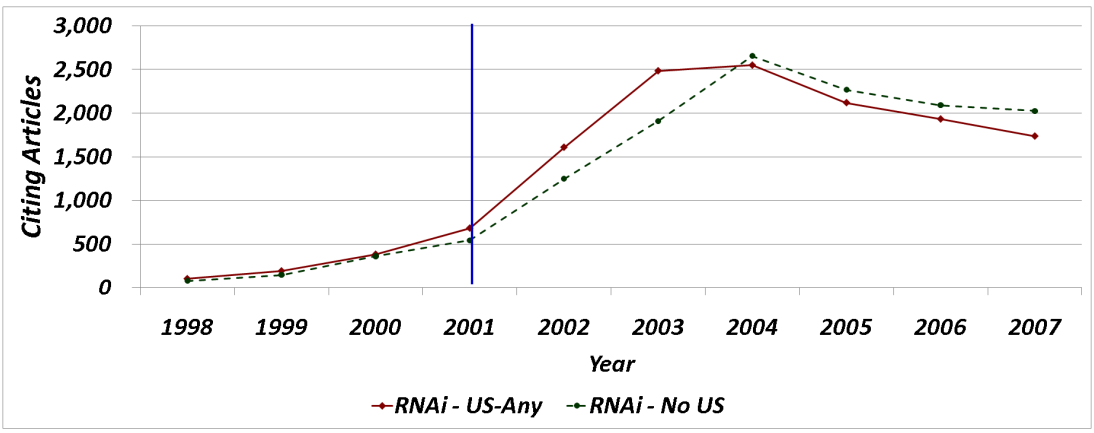
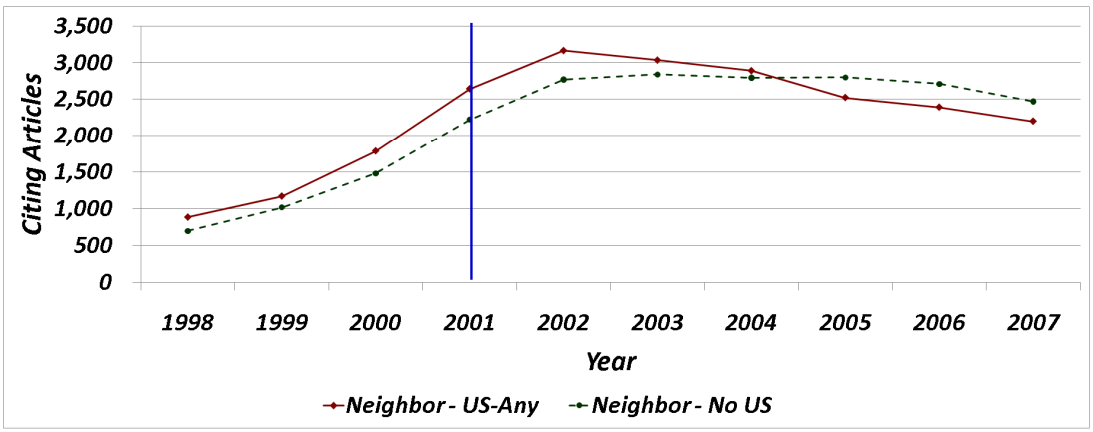


Figure #2 – Citing articles by publication type and year, US vs Non-US Reprint Authors









APPENDIX TABLE 1
US Stem Cell Research Policy in international comparison
(reproduced from Beardsley, 2005)

US

- Number of published hESC lines: 46
- Production of new lines: Legal, but prohibited with federal funds
- Therapeutic cloning: Legality varies from state to state
- Federal government funding: About $550m for all stem cell research ($24m for hESC)
- Private funding: About $200m
- Public funding at state level:
  California: $3bn over 10 years; New Jersey: $11.5m (another $380m proposed); Wisconsin: $375m proposed; Illinois: $1bn proposed; Connecticut: $20m proposed
- Federal government allows its funds to be used only on the 22 available hESC lines created before August 2001.

EU

- Production of new hESC lines: Permitted from unused IVF embryos where legal in member nations
- Therapeutic cloning: Prohibited
- Funding: $170m on stem cells over the past three years (only $650,000 for hESC research)
- Status in some member nations:
  - France: Creation of hESC lines from IVF embryos legal as of October 2004; public funding is $4m
  - Germany: Only work on hESC lines predating 2002 is legal; public funding is $4m
  - Finland: Permits research with IVF embryos; public funding is $5m
  - Italy: June 12 referendum will consider permitting IVF embryo research; public funding is $6m
  - ***EU will not increase funding for hESC projects despite a doubling of the total research budget.***

SWEDEN

- Number of published hESC lines: 8
- Production of new lines: Legal
- Therapeutic cloning: Legal as of April 2005
- Number of researchers: 400
- Government funding: $10m-$15m
- Private funding: Cellartis and NeuroNova, the two largest stem cell research companies in Sweden, contribute the bulk of the $35m spent annually there.
- Cellartis, the single largest source of defined hESC lines in the world, maintains more than 30--two of which are approved by the US National Institutes of Health.

UK

- Number of published hESC lines: 3
- Production of new lines: Legal
- Therapeutic cloning: Legal
- Government funding: About $80m
- Private funding: $15m-$20m:
  - The Wellcome Trust alone has spent $12m annually since 2002.
- First licence for human ES cell research was granted in 1996.
- The Human Fertilisation and Embryology Act of 1990 allows the UK to fund hESC research flexibly.
- UK's first licence for human cloning research granted in 2004. Its recipients announced in May 2005 the country's first cloned human embryo.

BRAZIL

- Production of new hESC lines: As of March, legal from IVF embryos at least 3 years old
- Therapeutic cloning: Banned
- Government funding: $4.5m annually planned, allocated by the Health Ministry and the Science and Technology Ministry

---

SOUTH KOREA

- Number of published hESC lines: 29
- Production of new lines: Permitted with case approval from Ministry of Health
- Therapeutic cloning: Permitted with case approval from Ministry of Health
- Number of researchers: 300-400
- Government funding: About $10m
- Private funding: About $50m

SINGAPORE

- Number of published hESC lines: 1
- Production of new lines: Legal, if embryos are destroyed within 14 days
- Therapeutic cloning: Legal, as above
- Number of researchers: About 150, in industrial and academic settings
- Academic spending: About $10m, from public and private sources
- Industrial spending: About $10 million
- A pending government proposal would spend $60m over the next four years.

ISRAEL

- Number of published hESC lines: 1
- Production of new lines: Legal
- Therapeutic cloning: Legal
- Government spending: About $5m
- Private spending: $15m-$30m
- ***Israeli scientists led one of the research teams that first isolated hES cells. They were also the first to show that hES cells could be changed into heart cells, and to show that hES cells can integrate with tissues.***

CHINA

- Production of new hESC lines: Legal
- Therapeutic cloning: Legal
- Number of researchers: 300-400
- Public and private funding: About $40m
- The journal *Nature* reports that "China has probably the most liberal environment for embryo research in the world", with little public opposition to such studies. No laws govern stem cell research, but the recommendations of the Ministry of Health endorse it.

AUSTRALIA

- Number of published hESC lines: 1
- Production of new lines: Conditionally legal
- Therapeutic cloning: Banned
- Number of researchers: 200-250
- Government funding: The Australian Stem Cell Centre has $90m to spend through 2011.
