Knowing What We're Getting:
Evaluating Scientific Research on the International Space Station

William Bianco

Eric Schmidt

Dept of Political Science
Indiana University

> ...it bothered me a little bit that I never saw in any scientific journal any results of anything that had ever come out of the experiments on the [space] shuttle that were supposed to be so important.
>
> - Feynman (1989, 72)

This paper brings evidence to bear on a central debate about the International Space Station: what, as Feynman asks, is the nation getting from its multi-billion dollar investment in station science? While NASA publicizes selected research findings and their alleged contributions to life on Earth (Robinson 2013), and describes virtually all activities onboard the ISS in terms of their scientific potential – an espresso machine was cited for its potential contributions to capillary fluid physics (Weislogel 2015) – neither the agency or anyone else has systematically measured or analyzed variation in the station's scientific output. This situation has persisted despite ongoing, often emotional debates within the scientific and political communities about the value of station science (Smith 2015), and the recent government report that criticized the lack of measurable targets to assess these activities (GAO 2015).

Our premise is simple: a proper evaluation of continued ISS funding requires a move beyond emotional appeals, vague generalities about the long-term value of space exploration, or assertions that the station contributes to peace between nations. We argue that station science can be evaluated using yardsticks that are familiar to researchers across scientific disciplines, such as the likelihood that research projects yield refereed publications, the quality of publication venues, and generation of patentable findings. To this end, we analyze the outputs of station science using a multivariate model that predicts the publications and patents from experiments as a function of time, project type, and affiliation of principal investigators. We find a relatively high probability that ISS experiments with PIs drawn from outside NASA will yield refereed publications. Furthermore, these experiments have significant probabilities of finding publication in high-impact journals or producing government patents. In this sense, ISS research satisfies a plausible first-order condition for good science: peer referees consider these results worthy of publication or patent protection. However, experiments whose primary purpose is technology demonstration – particularly if their PIs are all NASA-affiliated – have significantly poorer prospects for publication or patents.

These findings highlight the difficulty of constructing a compelling case for science onboard the ISS. The typical ISS experiment has a surprisingly good chance of producing findings that are worthy of a patent or publication in good journals. Even so, it will be years or even decades before these discoveries are refined into products or techniques that have a measurable impact on the lives of ordinary Americans – if indeed they ever do. In the meantime, NASA and other supporters of station science are left with the problem of justifying continued investment in ISS operations (totaling some 3-4 billion dollars a year for the US alone) given the lack of immediate, tangible benefits. While ISS operations have the disadvantage of requiring a large, highly visible budget, this problem is precisely the same as that faced by most practicing scientists, whose work is aimed at the generation of scientific knowledge rather than the development of commercially-viable products.

Lottery Tickets, Breakthroughs, and Normal Science

From the beginning of the International Space Station program, critics have highlighted the lack of a compelling research agenda (Leary 2002, GAO 2013). Contemporary accounts doubt the value of station funding (Matthews 2014), ask “where’s the science” (Matthews 2012), or wonder “what it is for” (Achenbach 2013). Even today, NASA’s metrics for describing station operations emphasize usage statistics, such as the percentage of laboratory space used, crew hours devoted to research, or the number of experiments (NASA OIG 2013). Moreover,

over a decade after the beginning of crewed operations (and two decades after the station program began), disagreements persist about which domains of natural science have the best prospects for producing high-quality findings (Launius 2015).

In truth, ISS experiments, like many other research programs, are lottery tickets. Even if a project has major scientific and commercial potential, advances may take years or decades to develop – if they ever do. Consider the research into microencapsulation, identified by NASA as a top ten research result: ISS experiments ended in 2002, with patents awarded to the PIs in 2009-2011, but clinical trials of drug delivery mechanisms were only in the planning stage as of 2013 (Robinson 2013). Development of commercially viable products is at least several years away, and may never happen. In this sense, station science resembles other big science projects (Weinberg 1961) such as the Large Hadron Collider, the James Webb Space Telescope, or the EU's Human Brain Project. Like ISS, these projects combine high up-front costs with uncertain, out-year payoffs (Fortin and Curry 2013). As a result, they are bedeviled with persistent criticisms about research goals and means.

More philosophically, research on station is dominated by normal science (Kuhn 1962) – "research firmly based upon one or more past scientific achievements, achievements that some particular scientific community acknowledges for a time as supplying the foundation for its further practice" (1962, 10). However mundane, the vast majority of scientific research takes the form of normal science, providing the basis for scientific breakthroughs when it generates results that contradict established theories (Fortin and Currie 2012). Still, there is no guarantee that a given project will produce anomalies, or that the scientific community will quickly appreciate whatever anomalies are produced (Dyson 2012). Thus, while scientists' ideal would be for ISS research to produce immediate and far-reaching scientific breakthroughs, normal science *by* *definition* guarantees that most experiments will not actually do so. Moreover, the full impact of even paradigm-shattering work may take years to develop.

Because of this time lag, the task of evaluating station science is similar to that faced in evaluating other big science projects: since the true value of a finding will only become clear over time, decision-makers faced with making a near-term judgment must rely on indirect indicators – refereed publications, assessments of publication venues, and patents (Fortin and Currie 2013). Publication in a refereed journal indicates that a project's methods and findings have satisfied peer review; publication in a journal regarded as influential implies that the findings were significant enough to survive an especially demanding review process. Patents signify that research has produced "a new and useful process, machine, manufacture, or composition of matter."[1] One high-impact ISS research result shows how the on-station zero-gravity environment affects the virulence of the *Salmonella* bacteria. While these findings have not produced any new products such as human vaccines, results were published in a high-Eigenfactor journal, PLOS ONE (Wilson et al 2008).

Productivity Data

Even if we opt to characterize station science in terms of publications and patents, our analysis confronts two difficulties. First, as shown in Figure 1, a significant percentage of experiments recorded by NASA are not scientific investigations.[2]

[Fig. 1. NASA-sponsored ISS Research, 2001-2014 – about here**]**

These data were assembled from NASA's web pages on station science, which provide information on about 900 separate experiments, of which 391 were listed as US-sponsored and

---

[1] Language from the United States Patent and Trademark Office website, http://www.uspto.gov/patents-getting-started/general-information-concerning-patents [accessed May 24, 2015]

[2] We use NASA's definition of an experiment as our unit of analysis, although in some cases, activities coded as separate experiments appear to be closely related work (i.e. follow-ons or variations of experimental conditions with the same research goals).

non-educational.[3] An additional 20 percent of so-called experiments were technology demonstrations or validations of station hardware which, given their purpose, appear to have a relatively low chance of producing new scientific knowledge. Our coding of the type of research project follows NASA's categories (Biological Sciences, Earth and Space Sciences, Human Physiology, Physical Sciences, and Technology Development) with one exception – we combine Earth and Space Science cases with Physical Sciences because of a limited number of cases in the former category.

Figure 2 shows current (August 2014) counts of refereed publications (RP), and high-value results (HVR) - publication in a journal with a five-year Eigenfactor score above the 90th percentile), and patents.[4] These data paint a dismal picture of station science, with over two-thirds of experiments failing to produce any output.

[Fig. 2. ISS Patents and Publications (as of August 2014) – about here**]**

However, this data does not reflect the surge in science activity after completion of ISS assembly in 2011. Figure 3 shows that nearly one-third of the experiments in NASA's database were either still grounded as of August 2014 or had been collecting data for less than three years – not enough time for data to be collected and analyzed, let alone for results to be peer reviewed.

[Fig. 3: Time and ISS Research – about here**]**

---

[3] We exclude educational projects, as they are clearly not aimed at producing publishable research. We exclude an additional five projects that are labeled by NASA as commercial demonstrations, such as filming cosmonauts playing golf during an EVA. We also exclude 11 projects that were conducted exclusively onboard space shuttles during ISS assembly missions, an additional 15 projects that were clearly educational (e.g. deployment of small satellites built by college students) yet listed as technology demonstrations, and 23 experiments that were listed in the database but had not yet been launched as of August 2014. With these restrictions (some of which overlap), the number of projects equals 337 (as reported in Figure 1). NASA's data on experiments is available at http://www.nasa.gov/mission_pages/station/research/experiments/ experiments_by_name.html [accessed October 5, 2015].

[4] NASA's description of each experiment separates publications into three categories – results, related, and ground-based. We focus on results publications only, as publications in this category are most directly a consequence of activities onboard the ISS.

Our multivariate analysis addresses these concerns by controlling for both the type of experiment and for the time since beginning data collection.

Multivariate Analysis

In our analysis, for each of the projects comprising figure 1, we determine whether the project produced RPs or HVRs and code the area being researched. We code the identity of Principal Investigators into three categories: all NASA PIs, all non-NASA PIs, and a combination of NASA and non-NASA PIs. We code time in months since the beginning of data collection using NASA's list of the ISS expeditions that conducted each experiment.

Using this data, we estimate two unordered logistic regression equations using the STATA logit procedure:

$$\text{RP} = \beta_{10} + \beta_{11}(\text{Time}) + \beta_{12}(\text{Time}^2) + \beta_{13}(\text{NASA}) + \beta_{14}(\text{Demo}) + \beta_{15}(\text{Bio}) + \beta_{16}(\text{Phy/ESS})$$

$$\text{HVR} = \beta_{20} + \beta_{21}(\text{Time}) + \beta_{22}(\text{Time}^2) + \beta_{23}(\text{NASA}) + \beta_{24}(\text{Demo}) + \beta_{25}(\text{Bio}) + \beta_{26}(\text{Phy/ESS})$$

Coding the binary endogenous variables is straightforward: RPs are coded 1 if the experiment generated refereed publication; and 0 otherwise. HVRs are coded 1 if the experiment either 1) yielded publication in a journal with an Eigenfactor score greater than the 90th percentile value in the 2013 Eigenfactor data, or 2) led to U.S. patent protection; and 0 otherwise.

The exogenous variables are Time (time in months since the beginning of the experiment's operations on ISS) and Time^2 (using the pair of linear and squared-transformation time variables is a standard technique to account for diminishing marginal returns from the impact of time on the dependent variables), NASA (1 if all of the PIs were NASA employees, 0 otherwise), Demo (1 if experiment was listed as a technology demonstration either in NASA's records or description of the experiment, 0 otherwise), Bio (1 if experiment listed as a biological sciences experiment by NASA, 0 otherwise), Phy/ESS (1 if listed as a physical sciences or earth and space sciences experiment by NASA, 0 otherwise). The omitted disciplinary category is thus experiments listed as human physiology. We expect that the Time variable will have a positive and statistically significant parameter; that the Time^2 parameter will be negative; and that the NASA and Demo parameters will be negative and significant. However, we have no firm expectations on the disciplinary variables.

This model specification allows us to control for the impact of time on the likelihood of different outputs from an experiment, while at the same time capturing variation across different types of experiments. Moreover, we can measure the impact of each exogenous variable at the margin of the others – for example, whether changes in observed patent rates are a function of time or a change in the mix of experiments conducted on the ISS. Finally, estimating the impact of time on publication (again, at the margin of the other variables) allows us to use the parameters to predict how publication rates will trend into the future.

Estimation Results

Table 1 gives the parameters derived from maximum likelihood estimation of this model, along with various measures of statistical significance and goodness-of-fit.

[**Table 1**: **Logistic Regression Parameters –** about here]

The parameter estimates are as expected. First, the likelihood of RPs and HVRs from ISS experiments increases over time. Second, there are no differences in the likelihood of RPs or HVRs across scientific disciplines. Finally, experiments with all-NASA PIs, as well as technology demonstrations, are less likely to generate RPs and HVRs.

Figure 4 illustrates these findings for two canonical cases: scientific experiments led by at least one non-NASA PI, and all-NASA technology demonstrations. Each plot uses our logit parameters to calculate the estimated likelihood that an experiment will yield an RP or an HVR as a function of the number of months since the beginning of data collection onboard ISS.

[Fig. 4: Predicting Outputs of Station Science: The Likelihood of Refereed Publications (RP) and High Value Results (HVR) – about here**]**

Clearly, the expected outputs from station science vary considerably over time. For example, two years after the start of data collection, the estimated likelihood that a scientific experiment yields an RP is only about .25. However, at 100 months, this probability has increased to almost .75. This variation confirms that station science is indeed a work in progress. Relatively low publication rates reflect the simple fact that many experiments have not been operating long enough to generate publishable results – let alone complete the journal or patent review process. Figure 4 also reveals a significant probability (about .50 after 100 months) that ISS science experiments will result in HVRs. Thus, while station science may not escape the doldrums of normal science, its findings are more likely than not to be published in the very best venues. Since these venues reflect a special imprimatur from the scientific community, research published there may well lead to scientific breakthroughs.

Not all experiments are created equal, however. Compared to scientific experiments with outside PIs, ISS technology demonstrations directed by all-NASA teams are significantly less likely to produce both HVRs and RPs. In one sense, this finding is not a surprise - by definition, technology demonstrations are unlikely to produce scientific results. Moreover, supervising PIs may lack the time and organizational incentives to draft their findings in a form suitable for publication or patent submission.

Discussion

Our results offer a general guide to the long-term research productivity of the International Space Station. While we cannot speak to the prospects of any particular project, many experiments are clearly producing noteworthy results. The rate of such outputs should improve significantly over time. Moreover, some areas of research onboard ISS are more likely than others to produce noteworthy results, with technology demonstrations lagging behind scientific experiments as a source of publications or high-value findings. These results highlight a stark tradeoff for future policy decisions. If ISS is primarily used to develop technologies for future crewed space exploration, this research will be less likely to produce journal-ready scientific advances. Thus, the very research most reflective of NASA's mission statement may be the hardest to sell to congressional overseers.

Even beyond technology demonstrations, though, ISS experiments are expensive normal science – combining high up-front costs with uncertain prospects of future benefits. Yet most practicing scientists (or review committees) would be highly satisfied if their projects yielded outputs similar to the typical ISS scientific experiment – a high probability of refereed publications, and a significant chance of high-value results. Such percentages are what first rate normal science looks like. Thus, to judge ISS research a failure because it has not yet produced breakthrough results is to ignore the nature of normal science; by this standard, virtually all research would be judged a failure. Of course, ISS experiments have the disadvantage of extraordinary cost and intense visibility; normal science is unromantic enough without the multi-billion dollar annual cost of station operations. It may be that getting out of ISS is a forward-thinking move on NASA's part. However, before abandoning the station, we should know what would be surrendered: normal science lottery tickets with uniquely high payout odds, and scientific returns certain to increase with each passing year.

References

Achenbach, J. 2013. “The skies. The limits. The international space station is one of humanity’s great engineering triumphs. But what is it for?” *The Washington Post*, September 14, 2013.

Dyson, F. 2012. “Is science mostly driven by ideas or by tools?” *Science* 338, 1426-7.

Feynman, R. 1988. *What Do You Care What Other People Think? Further Adventures of a Curious Character*, New York: W.W. Norton & Company.

Government Accountability Office. 2015. “International Space Station: Measurable Performance Targets and Documentation Needed to Better Assess Management of National Laboratory.” http://www.gao.gov/assets/670/669851.pdf, April 10, 2015, accessed August 1, 2015.

Fortin, J.M.; Currie, D.J. 2013. “Big science vs. little science: how scientific impact scales with funding.” *PLoS ONE* 8:e65263

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

Launius, R. 2010. “Why did the United States want to build a space station?” https://launiusr.wordpress.com/2010/12/03/why-did-the-united-states-want-to-build-a-space-station/, December 3, 2010, accessed May 24, 2015.

Leary, W. 2002. “Panel asserts reduced space station plans will limit scientific work.” *New York Times*, June 8*,* A1.

Matthews, M. 2014. “Critics doubt value of International Space Station science.” *Orlando Sentinel,* Jan. 24, A5.

Matthews, M. 2012. “International Space Station: critics ask, where’s the science?” *Orlando Sentinel*, May 12, A10.

NASA Office of Inspector General. 2013. “NASA’s efforts to maximize research on the International Space Station,” http://oig.nasa.gov/audits/reports/FY13/IG-13-019.pdf, accessed Feb. 28, 2015.

NASA Office of the Inspector General. 2014. “Extending the operational life of the International Space Station until 2024” http://oig.nasa.gov/audits/reports/FY14/IG-14-031.pdf, accessed Feb. 27, 2015.

Robinson, J. 2013. “Could you choose just one? Top International Space Station research results countdown.” http://www.nasa.gov/sites/default/files/files/Could-You-Choose-Just-One-2013-09.pdf, October 10, 2013, accessed Feb. 28, 2015.

Smith, M. 2015. “Bolden: If we don't pull together, we're not going to Mars.” *Space Policy Online*. http://www.spacepolicyonline.com/news/bolden-if-we-dont-pull-together-were-not-going-to-mars, May 8, 2015, accessed August 1, 2015.

Weislogel, M. 2015. “Space Station espresso cups: strong coffee yields stronger science. A Lab Aloft (International Space Station Research,” https://blogs.nasa.gov/ISS_Science_Blog/2015/05/01/space-station-espresso-cups-strong-coffee-yields-stronger-science, May 1 2015, accessed August 1, 2015.

Weinberg, A. 1961. “Impact of large-scale science on the United States.” *Science* 134, 161-164.

Wilson, J.W. *et al.* 2008. “Media ion composition controls regulatory and virulence response of Salmonella in spaceflight.” *PLoS ONE* 3, 392

Figure 1: NASA-sponsored ISS Research, 2001-2014

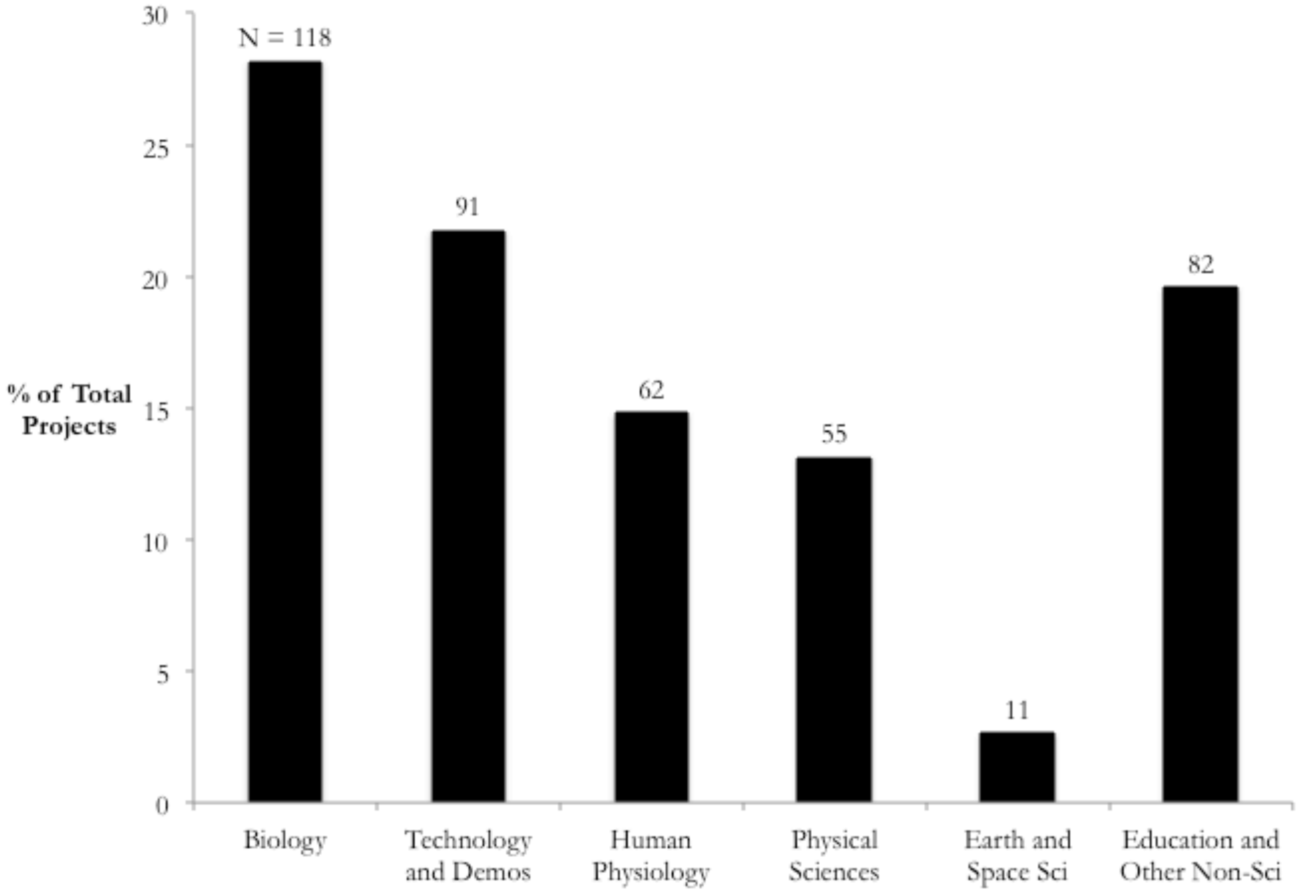

Figure 2: ISS Patents and Publications

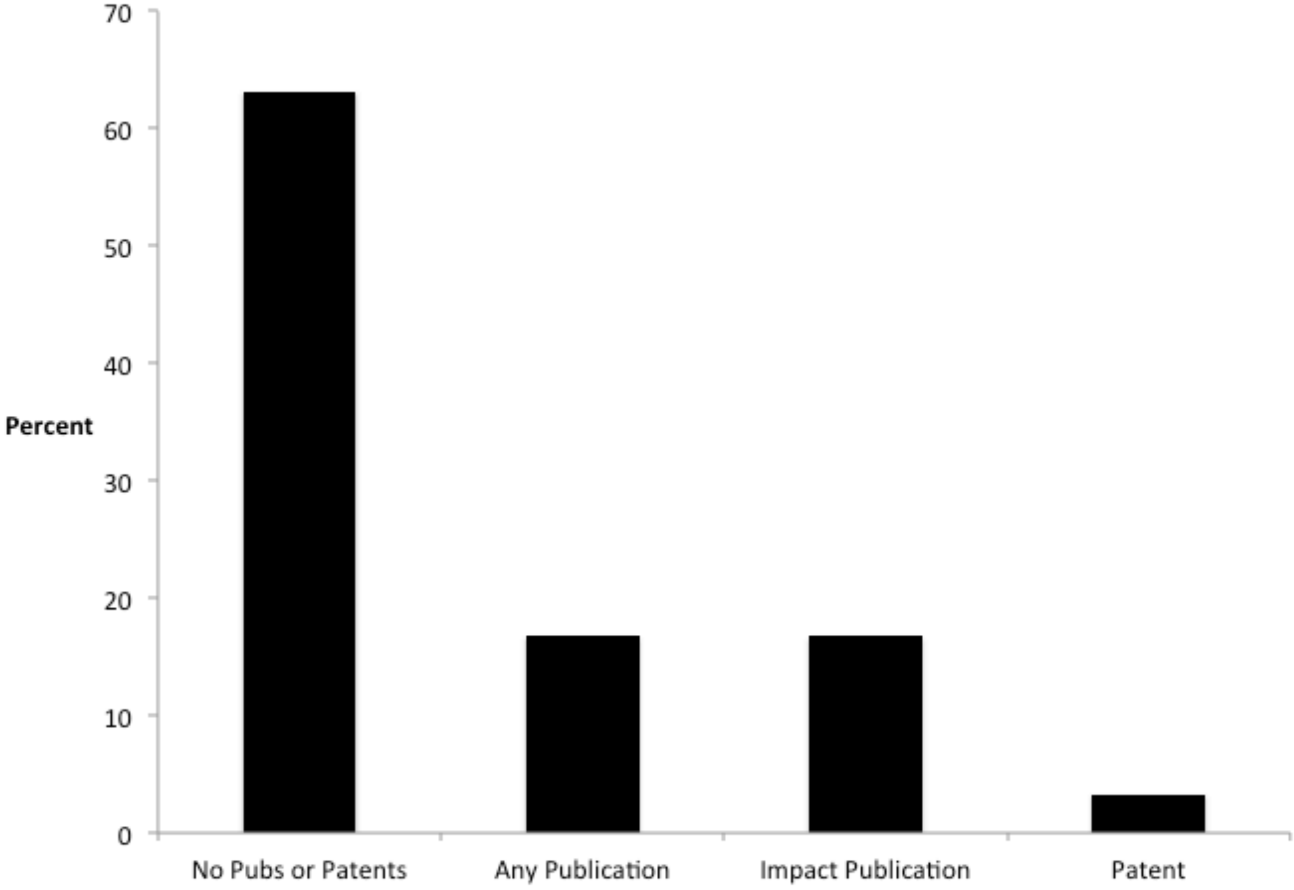

Figure 3: Time and ISS Research

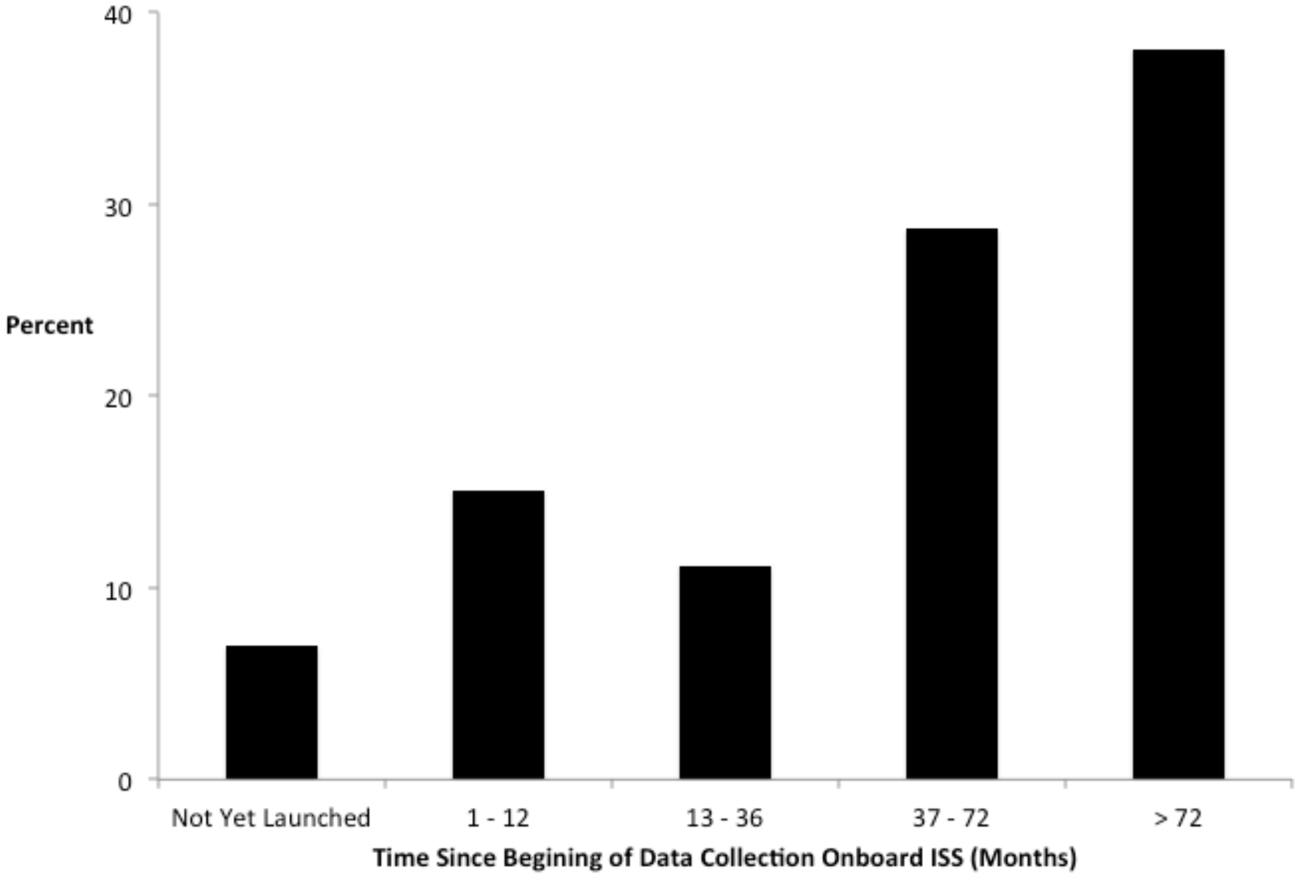

Figure 4. Predicting Outputs of Station Science:
The Likelihood of Refereed Publications (RP) and High Value Results (HVR)

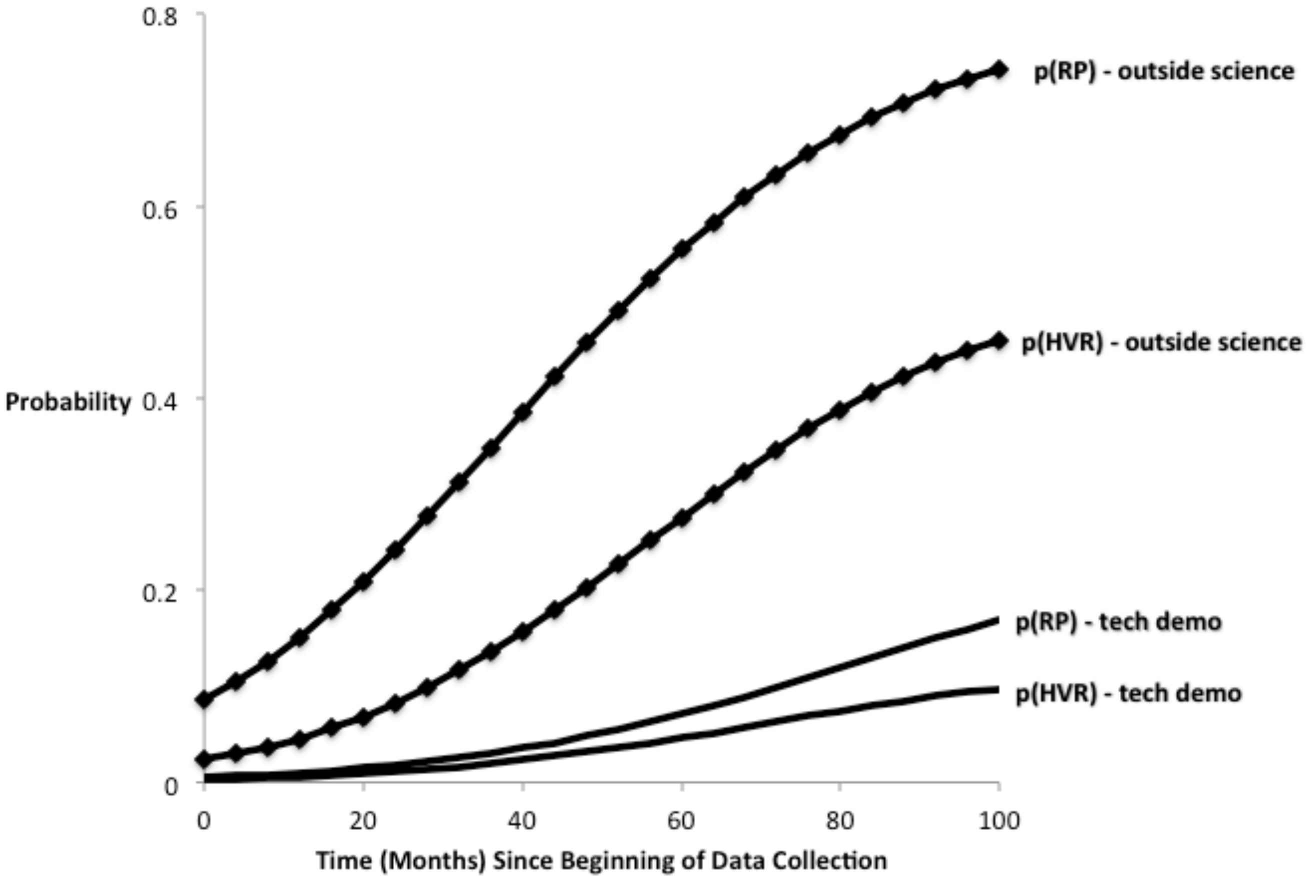

Table 1. Logistic Regression Parameters

Logistic Regression Parameters		
	Dependent Variables	
Independent Variables	RP	HVR
Time	.057*** (.011)	.062*** (.012)
Time^2	-.00022*** (.00067)	-.00026** (.00007)
Technology Demonstration	-1.89*** (.46)	-1.30*** (.52)
All-NASA PIs	-.99* (.55)	-.76* (.62)
Biological Sciences	-.41 (.36)	-.0044 (.39)
Phy/ESS Sciences	-.57 (.41)	-.13 (.44)
Constant	-2.37 (.49)	-3.76 (.53)
Model Chi Sq.	80.2***	56.4***

*** = p < .01, ** = p < .-5, * = p < .10, all one-tail.
N = 334 for both estimations.