What's in a (Missing) Name? Status Signals in Open Standards Development *

Timothy S. Simcoe
Boston University School of Management and NBER

Dave M. Waguespack
Robert H. Smith School of Business, University of Maryland

November 10, 2009

*We gratefully acknowledge financial support from the NET Institute (www.netinst.com) and the Kauffman Foundation. Lee Fleming was a catalyst for this research and provided valuable input on several early drafts. We also thank Ajay Agrawal, Scott Bradner, Brent Goldfarb and Olav Sorenson for helpful comments. Address for correspondence: Boston University School of Management, 595 Commonwealth Avenue, Boston, MA 02215. E-mail: tsimcoe@bu.edu

What's in a (Missing) Name? Status Signals in Open Standards Development

Abstract

How much are we influenced by an author's identity when evaluating their work? If names matter, are we responding to status, reputation or a signal of underlying quality? This paper addresses these questions in the context of open standards development. We exploit a natural experiment, whereby author names were occasionally replaced by *et al* in a series of email messages used to announce new submissions to the Internet Engineering Task Force (IETF), and compare the effect of obscuring high versus low status author names. Our results suggest that name-based signals can explain up to two-thirds of the difference in publication outcomes across status cohorts. However, this signaling effect disappears for a set of pre-screened proposals that receive more attention than a typical submission. We also show that high status authors receive more forward citations from other IETF participants than do low status authors, while cites from outside the focal community (from U.S. patents and academic journal articles) exhibit no difference. Our findings suggest that status signals are important for drawing attention to new ideas, which is important for developing these ideas and bringing them forward to publication. JEL Codes: L1, O3

1 Introduction

Hierarchy is a prominent feature of many, if not all, social systems (Gould, 2002). Sociologists have long emphasized the links between status, or position within a social hierarchy, and beliefs, actions or performance. Much of this research suggests that a higher slot in the social pecking order confers various benefits, such as increased deference, fewer constraints on behavior, undue credit and improved access to resources (e.g. Merton, 1968; Podolny, 1993).

Empirical research on status, however, typically comes with an important caveat: it is hard to establish causality. Observational studies must confront the possibility that status could be a function of performance, or that both status and performance might depend on some third factor, such as unobserved quality. While a laboratory experiment would solve this problem by artificially manipulating the social hierarchy (Ridgeway & Erickson, 2000), it is hard to find "natural experiments" that produce the same effect, especially when there are strong incentives for actors to seek the hypothesized benefits of higher status. Consequently, there is no evidence to date that can isolate the role of status effects in the field using archival data.

This paper takes a different approach to measuring status effects in the field. Instead of relying on variation in the social hierarchy, we focus on variation in the *signals* that actors use to learn one another's place. Specifically, we examine what happens when an author's identity is obscured. This idea of manipulating idnetity-based signals appears in Merton's seminal paper on status in scientific communities (Merton, 1968), which describes the experience of Lord Rayleigh, an eminent British mathematician and physicist who won the 1904 Nobel Prize in Physics:

> "Rayleigh's name was either omitted or accidentally detached [from a manuscript] and the Committee turned it down as the work of one of those curious persons called paradoxers. However, when the authorship was discovered, the paper was found to have merits after all."

Our study can be viewed as a large scale replication of the "Lord Rayleigh experiment" using data from the Internet Engineering Task Force (IETF), a community of engineers and computer scientists who develop protocols used to run the Internet.

The IETF has a review process to evaluate ideas for new protocols. At the start of this process, each new proposal is described in a short broadcast email announcement. Between 2000 and 2003, these announcements often used the generic label *et al* instead of providing a complete list of author names. This practice was driven by administrative overload at the rapidly growing IETF, particularly just before meetings, when staff were asked to process a large volume of new submissions. Because *et al* use is not correlated with author-level factors such as ability, reputation or influence, which might have a direct impact on publication outcomes, we can use name removals to estimate the causal effect of revealing an author's identity at the start of the IETF publication process.

To measure the impact of status signals, we compare changes in the probability of publication when high-status versus low-status names are removed from the email announcing a new proposal. As a proxy for status, we focus on authors that have served as an IETF Working Group chair; a position roughly equivalent to a journal editor. Our main finding is that replacing the name of a Working Group chair with *et al* leads to a larger decline in publication probabilities than obscuring the name of an average author. This finding is surprising for several reasons. First, the email announcements are only the first step in a lengthy and rather transparent community review process. Second, the *et al* "treatment" is much weaker than blind peer review; interested readers could easily track down missing names by examining the front page of the proposal itself. And third, the signaling effects are large, relative to a small overall chance of publication. Our point estimates suggest that name-based signaling explains as much as two-thirds of the benefits of having a high-status author.

After establishing that author names matter, we examine the underlying mechanisms and boundary conditions for this signaling effect. We begin by showing that name-based signals have no effect on publication rates for a set of proposals that are essentially "commissioned" by IETF Working Groups, and therefore have a much higher chance of publication. This finding is consistent with the view that status signals are more salient under conditions of high *ex ante* uncertainty about quality (Podolny, 1994).

Next we consider the role of *attention* as an intermediate input in the publication process. During the period we study, the IETF received about 80 new or revised technical manuscripts each week, averaging 23 pages apiece. Given the large volume of new ideas, identity-based signals may provide a simple heuristic for allocating attention to better proposals (Merton, 1968), or a useful cue for coordination (Morris & Shin, 2002). We show that proposals from high status authors generate more discussion on IETF email listservs, and go through more rounds of revision, when a high-status author's name is not obscured by *et al*. We interpret these results as evidence that status helps draw attention to a new idea, which is useful for developing the concept and bringing it forward to publication (Podolny, 2005, p.26).

Finally, we address the question of how name-based signaling effects should be interpreted. In particular, we ask whether screening on author names is based on a (possibly correct) belief that status is linked to latent idea quality, or alternatively, reflects the application of a double-standard that favors certain actors. While status is rewarded in either case, Merton (1968) calls the first mechanism "functional" since it ultimately selects for better outcomes. To distinguish between these theories, we draw from the economic literature on discrimination. Specifically, Becker (1993) notes that double standards, which economists call "taste based discrimination", subject the favored group to a less demanding review process. Thus, when screening is not functional, objectively measured *ex post* outcomes should be worse for the group that is favored *ex ante*. We implement this idea by comparing citation counts for publications by high and low status authors. To address the concern that identity also enters the citing process, we distinguish between citations coming from within the IETF and from external settings. Though our outcomes test has several shortcomings (discussed below), we find little evidence that high-status authors receive fewer cites, and tentatively conclude that name-based screening at the IETF is functional: author names convey information about the expected quality of proposals.

The main contribution of this paper is to exploit a unique organizational practice — the use of *et al* in a series of email announcements — to estimate the causal impact of status signals on publication outcomes in an open standards development community. Overall, our results suggests that an author's name provides an important *ex ante* signal, and that status is functional, since measures of status and quality are positively correlated *ex post*.

A second contribution is to highlight the role of attention as a resource that is disproportionately allocated to high status authors. Merton (1968) refers to this idea as the "Matthew Effect in the communication system." While we provide evidence for the attention mechanism in the novel organizational context of open standards development, we suspect that it generalizes to the larger scientific community, and perhaps other settings where attention is scarce and there is an important role for peer or community review (e.g. reviewing art, literature or cuisine). The idea that attention is an important intermediate input also helps to resolve the tension between theories that emphasize increasing returns to status, and those where screening is functional (and therefore fair) because status correlates with measures of objective quality.

Finally, this paper makes several methodological points that are particularly important for studies of status. First, we emphasize that in field settings, it will often be easier to to find random variation in signals than in the social hierarchy itself. Second, we illustrate the importance of finding an explicit source of random variation, since status signals are often highly correlated with other factors that may have a direct impact on performance. And third, we note a strong parallel between the sociological literature on status and the economic literature on discrimination. In particular, we re-introduce Becker's "outcomes test" as a tool for determining whether screening on status is functional or an example of taste-based discrimination. While our outcomes test is subject to a number of critiques, we know of no other paper in the literature on status that examines both *ex ante* screening and *ex post* outcomes to discriminate among alternative theories.

The remainder of the paper proceeds as follows: Section 2 reviews the literature on status and performance, with an emphasis on the role of status in scientific communities. Section 3 describes the Internet Engineering Task Force and the origins of the *et al* natural experiment. Section 4 describes our data, measures and statistical methods. Section 5 discusses the results, and Section 6 concludes.

2 Status Signals and Performance: Theory and Evidence

Sociological theories of status typically proceed in two steps, by arguing that (1) social positions and affiliations influence beliefs about a focal actor, and (2) beliefs effect the focal actor's performance (Podolny & Phillips, 1996). Thus, (Berger et al., 1980, p. 479) defines a status organizing process as, "any process in which evaluations of and beliefs about the characteristics of actors become the basis of observable inequalities."

This general framework can be applied to a variety of settings. For example, Merton (1968) describes how standing within the scientific community effects the distribution of resources (e.g. money, credit or attention) which matter for both output and recognition. In market settings, the link between status and perceived quality can lead to increased demand or bargaining power (Podolny, 1993; Benjamin & Podolny, 1999), and if legitimacy concerns induce low status actors to self-select out of some niches, high status actors may face less competition (Phillips & Zuckerman, 2001). These papers typically argue that status mechanisms are strongest in settings where there is considerable *ex ante* uncertainty about the underlying quality of a focal actor (or their products), since in that case others will rely more heavily on social affiliations, behavioral cues and other *status signals* when deciding how to act (Podolny, 1994).[1]

Building on these ideas, a large and rapidly growing literature provides evidence of the link between status and performance. For instance, Podolny (1993) measures status using a bank's relative position in securities underwriting advertisements, and shows that higher status investment banks have a greater spread between cost and price for a given offering. Benjamin & Podolny (1999) measure a winery's status in terms of network centrality (based on a matrix of cross-regional affiliations), and find that high status wineries charge more for bottles of equivalent quality. Washington & Zajac (2005), measuring status as a historical legacy of privileged treatment at the NCAA tournament, find that high status college basketball teams with marginal records are more likely to receive at-large invitations to the NCAA post-season basketball tournament. Stuart et al. (1999), measuring prominence based on centrality within

[1]Uncertainty about the true state of the world plays a similar role in economic models of signaling (Spence, 1973) and reputation Kreps & Wilson (1982). However, sociological theories of status drop the economic assumption that actors' beliefs after observing a signal are (on average) correct. For example, individuals may value high status activities or affiliations independently of their informational value (Gould, 2002); low quality actors may have the ability to mimic high-status traits; and signals that were once functional (the peacock's tail) may confer status long after changes in the environment have rendered the underlying information obsolete Feldman & March (1981).

networks of alliances and equity investors, find that biotechnology startups endorsed by high status organizations have initial public offerings faster and have greater initial public valuations. Stewart (2005) measures status as the number of third-party endorsements given to open source software developers, and finds that high status developers are more likely to receive a positive evaluation.

While this literature provides many insights into the link between status and performance, there are some basic questions that remain unresolved. For example, in a review of social capital research, Mouw (2006) notes that while theoretical distinctions have become increasingly refined, the field lacks solid evidence that there is a causal relationship between social capital of any form and outcomes, and is not even close to being capable of adjudicating among more refined alternative theories. We posit that finding evidence of status effects, and hence developing the theoretical literature, is hampered by three substantial empirical challenges.

2.1 Unobserved quality

At the root of theorizing about status is the argument that actor identity matters independently of intrinsic quality (Podolny & Stuart, 1995). However, it is very difficult to establish a *causal* link. The problem is typically described in terms of unobserved heterogeneity: some unmeasured dimension of quality might drive both status and performance. However, there is also the issue of reverse causation, where the hypothesized link from status to performance runs in the opposite direction, with variation on status simply an outcome of demonstrated ability (Gould, 2002).

Archival studies typically address this problem within a regression framework, by including a large number of controls for quality, power, ability or other factors that may have a direct impact on performance, and asking whether measures of status remain statistically significant. Many of these control variables are very clever (e.g. Benjamin & Podolny, 1999). However, if status and quality are highly correlated, this approach will always be sensitive to assumptions about measurement and functional form. Moreover, concerns over unobserved quality should be most salient in situations of great uncertainty — precisely the situations in which status is hypothesized as most relevant — because in these settings, the relative comprehension of the academic researcher is most diminished. In other words, in complex settings the academic researcher is himself a layperson, searching for social cues that are correlated with *ex ante* quality. It remains unknown whether true experts themselves rely on these social cues, or use other correlated indicators that were not included in the academic models.

One approach to this dilemma is to directly manipulate the status of a focal actor, holding all else constant. As a practical matter, this is often very difficult, particularly in field settings, since it violates modern norms of merit-based prestige. A second approach is to manipulate the signals that a third party might use to assess a focal actor's status. This approach is sometimes feasible, particularly when suppressing certain signals (e.g. race or gender) would promote merit-based evaluation (e.g. Blank, 1991; Goldin & Rouse, 2000). In principle, one could isolate the role of status by manipulating a very fine-grained set of signals: Who are an actor's friends? Where did they go to school? How do they dress? In practice, our signal is based on revealing an author's name — as in the Lord Rayleigh example — which might convey a great deal of information about status and other factors (Berger et al., 1980). Thus, while our experiment will not be able to disentangle all of the possibly salient attributes of a name, our estimates will place an upper bound on the causal impact of status-signals.

2.2 Increasing returns

Merton (1968) argued that scientific credit is often mis-allocated, with prominent scientists accruing ever greater recognition, while the contributions of less well known scientists are undervalued. A variety of mechanisms might explain why this doesn't result in a social system dominated by a single elite actor (Bothner & Haynes, 2009; Gould, 2002). Nonetheless, within these limiting parameters the implication of the status and performance literature is clear: greater status leads to competitive advantages, and hence greater status.

We propose that in the literature there are two distinct mechanisms offered for generating increasing returns to status. First is direct improvement in the status signal. In this conception, each success leads to better social positioning for the focal actor independently of changes in underlying quality. Recognition afforded to the focal actor becomes a cascade, with each act of deference increasing the odds that other alters imitate that behavior (Rao et al., 2001). For example, Merton (1968) describes how readers approach the work of a pre-eminent scholar with "special care" and in so doing, are apt to get more out of it.

The second explanation for increasing returns to status is intrinsic improvement in the quality of the focal actor's products. The source of these endogenous improvement may simply be behavioral: Merton (1968) notes that the validation received by a star scientist may give them the self confidence required to tackle important problems, while Phillips & Zuckerman (2001) posit that middle status actors are constrained by legitimacy concerns from taking risky actions. Endogenous improvements in quality may also result from the transfer of tangible resources. For instance, academic scientists with prestigious affiliations generally have greater resources at their disposal, and Sorenson & Waguespack (2006) find that film distribution companies invest more in marketing films from high status production teams. A more encompassing view is that actors with better social positions access more valuable knowledge, with improved management and outputs as a consequence (Powell et al., 1996; Burt, 2004). In this respect, the greater attention directed at high status actors may serve to increase intrinsic quality and not just lower advertising costs.

The improved signal explanation and improved underlying quality explanation for increasing returns to status are not mutually exclusive — both can operate at the same time. However, one can seek evidence of the increased underlying quality story by looking for observable differences in the behaviors of or resources used by high status actors, particularly during the *before* there is any evaluation of quality. We implement this idea by examining how the presence of a status signal influences, prior to the screening decision, the allocation of attention from third parties and the efforts of the focal actors themselves.

2.3 Efficiency

Even if status effects on screening decisions are demonstrated, the long run implications of that type of decision bias on efficiency are unclear. One school of thought emphasizes that status is functional, or at least loosely correlated with underlying quality (Podolny, 1993; Podolny & Phillips, 1996). When quality is difficult or expensive to measure, screening based on social cues can provide a reasonably efficient alternative. For instance, academic hiring committees may create short-lists based on individual and organizational affiliations due to the belief that these are quality indicators, and because it is very expensive to evaluate the portfolio of every applicant. This undoubtedly hurts some individuals, but if status is correlated *on average* with underlying quality, and search costs are high, then it is a useful tool for the screening actor.

Another school of thought emphasizes the "unearned" component of status, describing it as a construct possibly independent of underlying quality and thus essentially representing the application of privilege as opposed to a signal (Washington & Zajac, 2005). In this view status itself is the object of consumption (Burris, 2004). Moreover, having attained a privileged position, elites may become over-confident in their own abilities and therefore less likely to respond to external changes or develop new skills, ultimately resulting in worse performance (Bothner et al., 2009). We propose that these two views of status are hard to disentangle for *ex ante* screening decisions, but have very different implications for long-run *ex post* performance.

The economics literature on discrimination distinguishes between statistical discrimination, a social clue that is correlated with underlying quality, and taste-based discrimination, the application of double standards for privileged actors. Becker (1993) observes that taste-based discrimination implies worse *ex post* outcomes within the favored group because preferred actors face a lower *ex ante* selection threshold. Conversely, if status leads to better or no worse *ex post* outcomes, then this provides some evidence that status signals are a useful heuristic.

Unfortunately, readily comparable and objective *ex post* performance measures are often hard to find. Ayres & Waldfogel (1994) suggest that one such measure is market based pricing, a setting where non-functional discrimination is subject to competitive discipline. Another possible class of measure are head to head competitions, such as sporting events.[2] In order to carry out discrimination tests for other types of outcomes, the researcher must carry the assumption that social processes are at work in the screening stage, but not present in the long run performance indicator. For publishing decisions, citations provide a commonly used alternative quality metric. However, the question of whether cites reflect the *ex ante* quality of the focal document (it was better, therefore more subsequent work built on it), or other social processes, is a matter of debate. We propose that when using citations as a measure of *ex post* quality, a partial solution involves distinguishing between citations coming from within the community of practice, where citing is conflated with showing respect, and from outside the community, where citing parties have no stake in the relative prestige of the cited authors.

In the following section we discuss an empirical approach to understanding the relationship between status, quality, and performance that parallels the discussion above. To address the causality concern, we relay on a natural experiment — the replacement of author names by *et al* — and compare the effects of altering this signal for high and low-status actors. To address the issue of increasing returns, we examine the affect of status signal on the antecedents to the publication decision. To address the issue of whether status represents a signal or privilege, we examine future citations from both inside and outside the focal community to published drafts.

3 Organizational Setting

3.1 The IETF Publication Process

The IETF is a voluntary non-profit organization that creates and maintains compatibility standards used to run the Internet. Active participants are mostly engineers and computer scientists, representing a wide variety of academic, not-for profit and commercial organizations (Simcoe, 2008; Fleming & Waguespack, 2009). IETF standards are used to accomplish a wide variety of core networking functions, such as assigning IP addresses, routing packets and encrypting data. Firms that adhere to these protocols can reasonably expect their products to work with the rest of the Internet.

The IETF creates new standards and other practical networking knowledge using a bottom-up process of open publication and community review. Figure 1 provides an overview. Anyone

[2]For instance, a logical extension of Washington & Zajac (2005) would be to examine the effect of status on the performance, in the tournament, of basketball teams invited to participate in the NCAA tournament.

may propose a new idea to the IETF by creating an appropriately formatted document and submitting it as an Internet Draft (ID). Each new ID is posted to a public web server, where it remains for a period of six months. New IDs are debated and discussed on a series of email listservs maintained by the IETF and its various Working Groups, and at IETF plenary meetings, which are held three times per year.

The IETF distinguishes between two types of IDs: Working Group (WG) drafts and individual submissions. WG drafts are often "commissioned" as part of a broader technical agenda that is very likely to produce a standard.[3] While individual submissions may become WG drafts, or lead to the formation of a new WG, it is not very common.

Within six months of submission, one of three things will happen to an ID. First, the authors may decide to revise their proposal, usually in response to comments or concerns from the IETF community. Submitting a revision re-starts the ID's six-month publication clock.[4] Second, an ID may get published as a Request For Comments (RFC). And third, if an ID is not published, and its authors do not submit a revision, the draft will expire. Expired drafts are removed from the public ID repository.

Internet Drafts can follow two routes to publication. One possibility is to become the "consensus" recommendation of an IETF Working Group. If a Working Group draft has support from a solid majority of WG participants, the WG chair will forward the ID to the Internet Engineering Steering Group (IESG); a group of long-time IETF participants that act as a *de facto* editorial review board. For all WG drafts, the IESG issues a "last call" for comments to the entire IETF. If the last call raises serious issues, the ID is sent back to the Working Group for further consideration. Otherwise it is published.

The second way for an ID to become an RFC is through the "independent submission" process. In this case, authors submit their ID to the RFC editor, and ask that it be published as an individual (non Working Group) RFC. The RFC editor typically sends these drafts out to subject matter experts for review. If the reviewers and RFC editor agree that an individual ID is technically sound, and of general interest to the IETF community, it is sent to the IESG to ensure that it does not conflict with any Working Group drafts. If the IESG approves, the draft is published as an RFC.

On average, Working Group RFCs are more significant than independent submissions. For example, only WG drafts can be formally designated as IETF Standards. Individual submissions are published as "nonstandards-track" RFCs, and Simcoe (2008) finds that this leads to a

[3]In our data, Working Group IDs have a 47 percent publication rate, compared to 7 percent for individual submissions.

[4]Within the IETF, each revision is called an ID. For exposition, we call the entire series of linked publications an ID, and use the term revision to denote consecutive submissions.

much less contentious review process. Nevertheless, individual RFCs can be influential. They often propose new uses for IETF technology, describe lessons learned from implementation, and pose new problems for IETF members to work on.

3.2 IETF versus Academic Publishing

Perhaps because of the IETF's quasi-academic roots (Mowery & Simcoe, 2002), the IETF publication process is similar to academic publishing in several respects. Internet Drafts are formatted like academic papers. Manuscripts are authored by individuals. Published manuscripts typically undergo successive rounds of revisions, and an editorial board makes the final decision to publish. However, there are important differences between academic publishing and open standards development at the IETF.

First, while academic publishing typically uses blind reviewing, the IETF has an open review process, where all manuscripts are available online, and any individual may choose to offer feedback to the authors. Similar repositories of academic working papers, such as the Social Science Research Network (SSRN), are increasingly prominent in academic publishing, but do not themselves make determinations about which submissions are suitable for publication. For our research, one major advantage of the IETF's open review process is that outsiders can observe failed submissions, and not just successful publications.

A second difference between IETF and academic publishing involves the disposal of manuscripts. Submissions to academic journals are either accepted, rejected, invited to revise and resubmit, or withdrawn. At the IETF, the termination point for failed submissions is murky. According to IETF insiders, submissions rarely receive an "official" rejection. While authors may be strongly discouraged from pursuing a particular approach or idea, they always retain the right to continue revising an ID.[5]

Finally, the incentives and reward structure for IETF publishing is less clear than for academic publishing. While academics and career scientists must "publish or perish," IETF contributors are typically software engineers employed by firms. These engineers may contribute IDs as a way to advance their careers, or as part of their job when firms have a vested interest in particular standards. However, we have no direct evidence on how RFC publication influences individual outcomes, such as compensation or career mobility. Moreover, where academic publishing is usually characterized by a clear hierarchy among journals within a field, all contributions to the IETF are published as RFCs.[6]

[5]It is not unusual for an ID to go through 5 (or even 10) revisions before publication as an RFC. In a few cases, IDs will go through 30 or more revisions.

[6]The IETF does distinguish between standards-track and non-standards track RFCs, and among standards with different levels of maturity. These distinctions may play a role similar to that of the journal hierarchy.

3.3 The *et al* Experiment

Every ID submitted to the IETF is posted to a publicly accessible web page. With the initial posting, and for each subsequent revision, an e-mail announcement is sent to the "ietf-announce" listserv. The top panel in Figure 2 shows a typical ietf-announce message, containing a title, list of authors, filename (which contains information about WG affiliation and revision history), date, and abstract. The bottom panel in Figure 2 shows how the same information is presented in the actual Internet Draft: author names and affiliations appear on the front page, while detailed contact information is typically available at the back of the document.

Initially, every message sent to the ietf-announce listserv included the name of every ID author. However, beginning in 1999, some announcements replaced individual names with the generic label *et al*, or dropped them entirely. Figure 3 provides two examples.

Our empirical strategy exploits the fact that missing names sometimes belong to prominent members of the IETF community. In that case, *et al* removes a status signal from the ietf-announce message. Since interested readers could still find a complete list of names on the front page of the ID, the situation is akin to double-blind refereeing in the age of Working Papers and Google Scholar: *et al* introduces a certain amount of ambiguity, but does not make it especially hard to obtain the relevant information. In that sense, our paper is related to a literature on the effects of "salience" or attention costs (e.g. DellaVigna, 2008; Esteves-Sorenson, 2009).

The practice of using *et al* was introduced by the IETF Secretariat, an administrative body that manages the logistics of the ID publication process, to address a rapidly growing volume of submissions (see Figure 4). Once *et al* was allowed, the decision to use it for an individual ID was left to clerical staff, who would process incoming IDs by typing the relevant information into a standardized form. These individuals suggested that they tried to include every name, but would often resort to *et al* when things became busy — typically during the period just before IETF meetings, when there would be a spike in new proposals (Fuller, 2006).[7]

Figures 5 and 6 support the IETF Secretariat's explanation of *et al* usage. Specifically, Figure 5 shows a smoothed estimate of the probability that a draft with two or more authors receives an *et al*, where the vertical bars correspond to IETF meeting dates. There is a very strong cyclical pattern, with periodic spikes just before each meeting. Figure 6 shows that meeting-related deadlines also created large swings in the total volume of submissions over relatively short time-periods.

[7] For revisions, the clerical staff would typically cut and paste the original message into a new form. Thus, when *et al* appears on the message for an initial ID submission, it almost always remains for the entire life of the proposal.

4 Data and Methods

We collected data on all Internet Drafts submitted to the IETF between 1992 and 2004 from the website www.watersprings.org, and from the archived ietf-announce mailing list. The estimation sample includes all IDs submitted between 2000, when the *et al* experiment began, and 2003.[8] We drop single-author IDs because they never receive an *et al* on the ietf-announce message, and proposals with six or more authors, where the probability of receiving an *et al* approaches one.

For each ID, we use public sources to obtain the submission date, a complete list of author names (from the ID text), the number of revisions and an outcome: expiration (failure) or publication as an RFC. For each author, we have a complete list of their submissions to the IETF, their place in the list of draft authors, an email address, and information about whether they ever served as a Working Group chair. Table 1 provides a list of variables and definitions, while Table 2 provides summary statistics for the 3,719 individual IDs and 1,326 Working Group IDs in our estimation sample.

4.1 Measurement

The top half of Table 2 focuses on outcome measures. Our primary outcome is the indicator variable *Published as RFC*. Unconditional publication rates are 6.9 percent for individual IDs and 46.8 percent for Working Group IDs. We also have several intermediate outcomes. *Email Lists* is a count of the email lists where a focal ID is discussed. *Replies* counts the number of "reply" messages that mention an ID.[9] Both variables were constructed by an automatic search through the IETF email archives for any string that matched an ID's unique file name. Our last intermediate outcome is *Revisions*, which counts the number of times an ID was submitted. Not surprisingly, *Revisions* is highly correlated with publication. There are substantial differences in the means of all three intermediate outcomes for individual and Working Group IDs, suggesting that WG drafts receive more attention than individual submissions.

The last three rows in the top half of Table 2 are based on citations to published RFCs (note the large drop in sample size). *RFC Cites* was constructed by an automated search of the reference section in all published RFCs. Data on *Patent Cites* were collected via the US Patent and Trademark Office web site. *Article Cites* uses data from the ISI Web of Science database. The latter two variables count all citations that match the string "RFC" or "Request

[8] We drop IDs from 2004 because the IETF adopted a set of procedures for using *et al*, and to avoid truncation of the dependent variable, since some IDs take several years to be published.

[9] *Replies* is a better measure of "conversation intensity" than total messages because it excludes a variety of administrative emails that may mention an ID.

for Comments" followed by a four digit number.

The bottom half of Table 2 provides summary statistics for explanatory and control variables. The indicator variable *Et Al Dummy* equals one if the ietf-announce message for an ID either lists some authors as *et al*, or simply omits their names.[10] Twenty percent of individual submissions and 14 percent of WG drafts receive an *et al*.

Our primary measure of status is the dummy variable *WG Chair Author*, which equals one if an ID author has ever served as a Working Group chair. When an ID has a WG chair author, but the name of that author is not visible on the ietf-announce message — so there is no status signal — we set the dummy variable *Unlisted WG Chair* to one. Note that the mean of *Unlisted WG Chair* (3.4 percent) is smaller than the product of *Et Al Dummy* (19.5 percent) and *WG Chair Author* (25.9 percent), since some ietf-announce messages may contain an *et al* that only obscures the name of non-chair authors. Given the small incidence of *Unlisted WG Chair*, our main results rely heavily on outcomes in a sub-sample of 127 proposals where the name of a WG chair was obscured or omitted.

As a proxy for status, *WG Chair Author* has strengths and weaknesses. The main strength is that it is easy to interpret and almost certainly correlated with "true" status within the IETF community. To become a chair, individuals must be put forward by a WG, and approved by an Area Director who sits on the IESG. A chair's job combines elements of journal editor and parliamentarian. Specifically, this individual manages the publication process for all ID's associated with a WG, and decides when the group has reached consensus on substantive issues. Chairs have high visibility within the IETF, and most have done some piece of work that is viewed as significant. The main weakness of *WG Chair Author* is that it is a very crude measure of status. As a dummy variable, it is not capable of capturing subtle variations in the structure of the IETF's social hierarchy.

To highlight the close connection between status and past performance, we collected data on each author's publication history within the IETF. Thus, the variable *Published RFCs* provides a second proxy for status, based on the notion of a merit-based prestige hierarchy. Once again, we use the *et al* experiment to separate signaling effects from the direct impact of skill, knowledge, ability or experience. Specifically, the variable *Unlisted RFCs* measures the (missing) status signal by counting past publications of authors who contribute to the ID, but do not have their name listed on the ietf-announce message.

While there is no mechanical link between publishing RFCs and becoming a chair, Fleming & Waguespack (2009) show that there is a strong correlation between the two measures, and we find a similar relationship in our data at the ID level (see Appendix Table A1). This strong

[10] In robustness tests, we find little difference between omitting names, and obscuring them with *et al*.

correlation highlights a limitation of our analysis. While the *et al* experiment can be used to isolate the effects of name-based signaling, or what Merton (1968) called the Matthew Effect in the communication system, it will not reveal how readers interpret these signals. Thus, a high-status name may signal the expected quality of the proposal, the future importance of the idea (or technology), the political clout of the author, or the likelihood that others will pay attention to the ID.

The remaining rows in Table 2 provide descriptive statistics for the control variables. An average ID in our estimation sample has between two and three authors, and is roughly 11 pages long. The means of *Published Author* show that 43 percent of individual IDs, and 77 percent of Working Group drafts, have at least one author who has previously published an RFC. The other controls include a dummy for authors from outside the United States (*Intl Author*), and a dummy for ID's whose authors have more than one primary affiliation (*Multi-sponsor*). We use *Days-to-meeting* to control for "congestion effects" around meeting dates, as described by the IETF clerical staff.

4.2 Methods

4.2.1 Signaling Models

Our statistical methods examine the change in publication rates when a high-status author's name is removed from the ietf-announce message advertising a new ID. A simple way to measure the impact of status signals would be to ignore IDs from low-status authors, and use a simple t-test to compare the average publication rates when a WG chair's name is obscured or not; however, this approach will over-estimate the signaling effect if other factors produce a negative correlation between *et al* and publication. We are particularly concerned that meeting-related "congestion effects" (see Figures 5 and 6) may cause both an increase in *et al* usage, and a decline in publication rates. For example, proposals submitted during the pre-meeting rush may receive less attention, or be less "polished" than comparable proposals submitted at an earlier date.

To control for omitted variables correlated with *et al*, we compare the change in publication rates when *et al* obscures high versus low-status author names. Put differently, we use IDs where *WG Chair Author* equals zero to estimate the baseline impact of having an *et al* on the ietf-announce message, and subtract this baseline from the change in publication rates when *et al* obscures a high-status name. In this approach, the status-signaling effect is estimated by a difference in differences: the change in publication rate associated with removing a high-status author's name, minus the the change in publication rate associated with removing a low-status author's name. We implement this idea using the following linear regression

$$RFC_i = UnlistedChair_i\beta_1 + WGChairAuthor_i\beta_2 + EtAl_i\beta_3 + \lambda_t + X_i\delta + \varepsilon_i \quad (1)$$

where i indexes the Internet Drafts in our estimation sample; X_i is a vector of control variables; and λ_t are a vector of tri-annual IETF meeting dummies that control for time-trends in the underlying probability of publication. Though our main outcome is bounded at zero and one, we estimate a linear probability model for two reasons. First, it is easy to interpret coefficients as a change in probabilities. Second, and perhaps more importantly, in nonlinear models the interaction term (β_1) will not equal the cross-partial derivative of the expected probability of publication (Ai & Norton, 2003).

In equation (1), β_1 measures the signaling effect. The change in probabilities associated with having a high-status author is β_2 (though this coefficient may reflect a variety of other factors, such as latent proposal quality or author-experience). Ideally, we would let β_1 vary according to the position of the WG Chair in the list of authors, since an author's place may be determined by contribution to the ID, which may in turn influence quality. In practice, we do not have enough data to generate precise estimates in such a model, though we do explore this extension in the appendix.

After estimating the impact of name-based signals on publication outcomes, we use the same model, based on equation (1), to examine several intermediate outcomes. We interpret these outcomes as measures of the attention received by an ID during the review process. A draft that receives very little attention may be more likely to be abandoned for a number of reasons: attention is needed for a Working Group to become interested in a proposal; authors may interpret lack of interest as an assessment of quality; and the quality of drafts that do receive early feedback may improve as a result. Conditional on having an interesting idea, there may be a large difference between receiving a little attention or none at all, thus helping to explain how our "weak" treatment produces a measurable impact on publication rates. Given the rapid increase in total proposals, attention from prominent IETF members also seems likely to have been a scarce resource during our sample period.

4.2.2 Outcome Tests

Our final set of analyses will ask whether status-based screening at the IETF is functional, as opposed to merely an expression of preference for high-status authors. For this, we turn to a different method. Becker (1993) proposed that *ex post* outcomes might be used to distinguish between "taste based" discrimination, where signals convey no information about expected quality, and statistical discrimination based on traits that provide an accurate signal of performance. Ayres (2002) provides an overview of the approach.

In our setting, a natural measure of *ex post* performance is the number of citations received by an RFC. Intuitively, if WG chairs get a "free pass" in the review process, a chair-authored ID that just manages to reach RFC should receive fewer cites than a non-chair's barely accepted RFC. To be specific, suppose that cites are a valid measure of *ex post* quality, and that the publication process generates a set of quality thresholds c_s, where $s = H, L$ identifies an author's status. An unbiased screening process would generate a common set of quality thresholds $c_H = c_L$, so the distribution of cites conditional on publication is identical across status-groups. Taste-based discrimination in favor of high-status authors would lead to a lower quality-threshold for that group, so that $E[Cites|s = H] = E[Cites|q > c_H] \leq E[Cites|q > c_L] = E[Cites|s = L]$. We apply this idea within a regression framework, using $ChairAuthor$ as a measure of status. Specifically, we use Poisson regression to estimate the following model, where i indexes RFCs published in year t.

$$E[Cites_i] = \exp\{\alpha ChairAuthor_i + \beta X_i + \lambda_t\} \tag{2}$$

Becker's insight is that if cites are an unbiased measure of *ex post* performance, then systematic discrimination in favor of Chairs would lead to $\alpha < 0$.

The outcomes test has been criticized for several reasons. First, it makes predictions about the quality of a *marginal* as opposed to an *average* proposal. Thus, in order to operationalize the outcomes test, we must assume that proposals from high and low-status actors are drawn from the same underlying quality distribution. For example, if high-status authors are more likely to produce truly exceptional RFCs, an outcomes test may find mean citation rates are equal, even if there is discrimination against the low status types.[11]

A second criticism of outcome tests is that it can be hard to determine whether an outcome measure is itself contaminated by discrimination.[12] One approach to this problem is to use prices (as in Ayres & Waldfogel, 1994), or some other *competitive* performance measure, on the assumption that competitive outcomes are unbiased. Lacking an obvious market-based performance measure, we consider an alternative approach. Specifically, we compare forward citations from three different sources: other RFCs, US patents and academic publications. We expect academic publications and especially patents to be less biased, since they are produced by individuals who are (on average) less involved in the IETF, and in the case of patents, have no

[11] Knowles, Persico & Todd (2001) developed a formal model of discrimination in a search process which generates the prediction (2) even when there is unobserved heterogeneity — including differences in the quality distribution across groups. The key to their model (and the resulting "hit rate" test) is that high and low status authors respond to the incentives created by the search process when deciding whether to submit a proposal.

[12] For example, in the literature on discrimination in police searches, most authors use convictions as the outcome. Of course, this is problematic if the judicial system is prejudiced against certain defendants.

incentive to draw attention to prior work. While the resulting multiple-outcomes test alleviates some of the obvious concern for bias in the citation process, it also raises questions about the existence of a one-dimensional quality measure. In particular, the objective function that IETF members apply to the screening process may not coincide with the quality perceptions of academic paper-writers or patenters. Nevertheless, the use of multiple performance measures brings additional information to bear on the question of whether the signaling effect is a case of statistical or taste-based discrimination, given the possibility of similar biases in the citation process.

At the end of the day, we remain cautious about the results of the outcome tests. However, we believe they are worth including for two reasons. First, the results offer some additional support for our idea that the key mechanism in this setting is attention. Second, we believe outcome tests are a methodology that has been overlooked in the recent empirical literature on status. In particular, we know of no other study that combines direct tests for the impact of status (or status signals) with outcome tests that may yield insights into the underlying mechanism; specifically, whether it is functional or purely discriminatory.

5 Results

Tables 3 through 5 present our main results. Tables 3 shows that identity-based signals influence publication outcomes for individual submissions, but not Working Group IDs. Table 4 focuses on intermediate outcomes, and suggests that status signals influence publication outcomes for individual IDs by drawing more attention to proposals during the review process. Table 5 presents the citation-based outcome tests, which provide some preliminary evidence that name-based screening at the IETF is functional.

5.1 Identity as a Signal

Table 3 shows that name-based signaling has a significant impact on individual ID publication rates, where signals account for roughly two-thirds of the total benefits of high-status authorship. However, name-based signals have no impact on the publication of Working Group IDs. All of the results are based on the linear probability model described by equation (1).

The first column in Table 3 presents our simplest specification, which includes a constant, *WG Chair Author*, *Et Al Dummy*, and the key status-signaling variable *Unlisted Chair*.[13] The results show that individual IDs are 10 percent more likely to be published when the name of a WG chair appears on the ietf-announce message. However, this chair-author effect declines

[13] In this specification, there is no possibility that predicted values fall outside the unit interval.

by 6.6 percentage points when the author's name is obscured or omitted. Thus, name-based signaling account for 66 percent of the total benefits of WG chair authorship. Both the main effect and the signaling effect are large in comparison to the 5.5 percent baseline publication rate for IDs with no chair and no *et al.* Finally, our estimates show that receiving an *et al* is associated with a 5 percent drop in publication rates for IDs without a chair author. We interpret the negative main effect of *et al* as the impact of congestion, produced by the large influx of new proposals over the entire sample period, and particularly just before IETF meetings. This negative coefficient on *Et Al Dummy* implies that including the low-status control sample reduces our estimate of the signaling effect.

The second model in Table 3 shows that we find similar effects using *Published RFCs* as a proxy for status and *Unlisted RFCs* to measure the (missing) signal. This specification also includes a complete set of meeting effects, to capture time-trends, and dummies for the number of authors (between two and five). Once again, the results indicate that name-based signaling accounts for roughly two-thirds of the benefits from having a high status author. In this model, the main effect of *et al* declines substantially relative to the 8.4 percent baseline publication rate, perhaps because the IETF meeting effects capture time-trends in the overall publication rate that are correlated with an increasing number of *et al*'s.

The third column in Table 3 includes both proxies for status and a full set of controls. While the *Unlisted RFC* coefficient remains negative and significant, the coefficient on *Unlisted WG Chair* is essentially zero. It is not surprising that the signaling effects are stronger for the quality-based status measure, since RFC counts are highly correlated with having a WG chair author, but provide more variation than a simple dummy variable.[14] A test of the joint null hypothesis that both signaling coefficients equal zero is strongly rejected ($p = 0.00$). Still, these results highlight the main limitation of our methodology: it is not possible to un-bundle the many pieces of information potentially conveyed by a single name. Nevertheless, the *et al* experiment does allow us to separate the impact of an identity-based signal from variation in status itself, which appears to be highly correlated with one's past performance. Finally, note that the main effect of *et al* is no longer statistically significant in this specification, which includes the variable *Days-to-meeting* as a control for short-term congestion effects.

Since the results in the first three columns of Table 3 report our main finding, we subject them to a variety of robustness checks (presented in Appendix Table A2). We find very similar results, albeit with less precision, when the *Unlisted Chair* coefficient is allowed to vary according to the chair's position in the list of authors. This is reassuring, since one might worry about a correlation between *Unlisted Chair* and latent quality if chair authors are only obscured when

[14]See Table A1 for evidence of the strong ID-level correlation between RFC counts and chair-authors.

they make a smaller contribution to the underlying idea. The main results are also robust to dropping all proposals with a chair in the first-author position. We find no difference in either the main effect of *et al*, or the signaling effect, when an author's name is obscured by *et al* as opposed to inadvertently omitted from the ietf-announce message. Finally, we show that the *et al* effect increases substantially if we omit the low-status controls, as suggested above.

Returning to Table 3, the last two columns examine the impact of status signals in the sample of Working Group drafts, where there is much less uncertainty about proposal-quality and a higher chance of publication. Column 4, returns to the very simple specification used in the first column. While we find a much larger effect on *WG Chair Author*, the point estimate on the status-signaling parameter is almost precisely zero. This is consistent with the idea that status effect are more pronounced in environments characterized bey high uncertainty. The results in column 4 also make an important a methodological point. In particular, there is a large, positive and statistically significant coefficient on our proxy for status, eve though we conclude that there is in fact no signaling. The final set of results shows that we find the same pattern after introducing a full set of controls. As with the individual IDs, the main effect of *WG Chair Author* drops substantially, while *Published RFCs* enters as positive and highly significant. We cannot reject the hypothesis that both signaling coefficients equal zero ($p = 0.54$).

We draw two main lessons from Table 3. First, based on the individual ID results, we conclude that identity based signaling does play a role in the IETF publication process. Remarkably, our estimates suggest that signaling explains roughly two-thirds of the total effect of having a high-status author on a proposal. Second, comparing the individual ID results to the those in the Working Group sample, we conclude that this signaling effect is absent when proposals are viewed as more important, and thus have a higher *ex ante* chance of publication.

This second lesson suggests an interesting twist on the widely held view that status signals are more salient in an environment of high *ex ante* uncertainty. In particular, there is no exogenous difference between individual and Working Group IDs that suggests greater uncertainty in one case or the other. Rather, the level of uncertainty depends on the decision of individual IETF participants, and the entire community, to pay more attention to a set of particularly important initiatives. Thus, differences in *ex ante* uncertainty will often reflect collective judgements, and the expected costs of applying alternative search mechanisms or screening heuristics to learn about more salient attributes of the actor (or idea) in question. In practical terms, status is used as a signal to screen individual IDs precisely because of their low success rate, which makes it too costly to devote much attention to all of the new ideas being submitted. Our next set of results explore this hypothesis more systematically by focusing on proxies for attention as a intermediate input in the publication process.

5.2 Status and Attention

Table 4 uses the same empirical model, based on the *et al* experiment, to examine a set of intermediate outcomes; specifically *Email Lists*, *Replies* and *Revisions*. Collectively, we interpret these measures as indicators of the amount of attention that an ID receives after it is released. We focus on *WG Chair Author* as a proxy for status to simplify exposition, but show in Appendix Table A3 that we obtain similar results when *et al* is interacted with *Published RFCs*.

The first two columns in Table 4 focus on *Email Lists*. The results show that when status-signals are removed from an ietf-announce message, an ID is discussed on fewer listservs. Once again, signaling accounts for roughly two-thirds of the main effect of *WG Chair Author* in our baseline specification with no controls. When controls are introduced, the main effect of having a chair author falls, but there is little change in the size of the signaling effect.

The middle columns in Table 4 repeat this exercise using *Replies* as the outcome variable. We find that IDs from high-status authors generate more email replies than an average ID. The status signaling coefficient is negative and large (relative to the main effect of *WG Chair Author*), but not statistically significant. The last two columns in Table 4 focus on *Revisions*. This outcome is closely linked to final publication, since most unpublished IDs are abandoned after a small number of revisions, while publication often requires several rounds of editing. Once again, we find a statistically significant main effect of *WG Chair Author*, and the coefficient on *Unlisted Chair* indicates that a substantial component is linked to name-based signaling via the ietf-announce listserv. Interestingly, if we use *Published RFCs* as a proxy for status (see Table A3), the signaling effect on *Revisions* becomes statistically significant, while significance is lost on *Replies*.

Overall, the results in Table 4 indicate that high-status authors receive more attention, and that much of this effect is caused by placing their name on the ietf-announce message. While this attention effect is plausible — particularly given the large volume of IDs and preliminary nature of many proposals — it is nevertheless surprising that a difference of one or two email lists (or messages) can lead to a substantial divergence in publication probabilities. We interpret these findings as evidence of strong increasing returns to attention in the early stages of the creative process. It is unclear whether this is driven by unique features of the IETF publication process, or is a more general feature of creative work. However, we believe this to be an interesting subject for future research.

5.3 Status and Citation

Our final set of analyses ask whether status-based screening at the IETF is functional. Put differently, we test the idea that high-status authors get a "free pass" in the publication process against the null hypothesis that they produce better ideas, which would add an element of efficiency to status-based screening. Our approach is based on the idea that "taste based" discrimination (i.e. a free pass for high-status authors), should lead to worse *ex post* performance. While citations are a natural performance measure, they might also reflect a bias for high-status authors. We address this concern by examining cites from three different sources: RFCs, U.S. patents and academic journals. If the relative citation rate of high-status authors is smaller outside the IETF community, we might reasonably conclude that citations are themselves reflective of status.

We implement these citation-based outcome tests by estimating equation (2) using a Poisson regression, with robust standard errors to correct for over-dispersion. For this analysis, we expand the estimation sample to all individual IDs submitted between 1995 and 2003 that were published as RFCs. While this leads us to include a number of IDs from outside the period of the *et al* experiment, it also produces a significant increase in the number of RFCs in the estimation sample. We assume that citations are not influenced by (missing) status signals on the ietf-announce listserv, since RFCs are published with a complete list of names.[15]

The first two columns in Table 5 examine forward-citations from RFCs. The first column presents results based on a simple specification that only includes controls for RFC publication year (because of truncation in the citations data) and the number of authors. We find that there is a significant positive correlation between *WG Chair Author* and RFC citations. This strongly rejects the prediction of "taste based" discrimination, and even overshoots the null hypothesis that status-signals are (on average) correct. One interpretation of this result is that RFC citations are themselves biased in favor of high-status authors; perhaps even more than the screening process. A second possibility is that status-based screening is functional, but the skewed nature of the citation distribution leads to poor performance of outcome tests based on the assumption that high and low-status RFCs have an identical infra-marginal distribution of citation counts. The second column in Table 5 adds a variety of controls, including a count of all authors' future RFCs, which provides a rough proxy for self-citations, and finds a similar positive correlation between past publications and future cites.

The next two columns in Table 5 examine cites from U.S. patents. In this case, we find no evidence of a difference in the citation rate of high and low status authors. While the point

[15]As a practical matter, dropping RFCs where there was an *et al* makes little difference for the results, since they are a small proportion of the total sample.

estimate on *WG Chair Author* is positive, it is not very precise. Adding controls does little to clarify the picture. Since we cannot reject the null hypothesis of no difference, these results suggest that name-based screening is functional. However, to reconcile this conclusion with the large positive coefficients in the first two columns, we must also believe that there is an element of taste based discrimination in the citing practices of authors within the IETF community.

The last two columns in Table 5 examine citations from academic journals. In this case, the simple specification produces a negative coefficient on the *WG Chair Author* coefficient. While this is consistent with the idea of discrimination in favor of WG chairs, the parameter estimate is imprecise, and we cannot reject a null hypothesis of zero. Adding controls causes the coefficient on *WG Chair Author* to become negative and significant, while the corresponding coefficient on *Published RFCs* is positive. While a test of joint significance rejects the null hypothesis that both coefficients are zero (p=0.04), neither enters this regression significantly if the other is excluded. Thus, while these results hint at taste-based discrimination, we view them as inconclusive. Since no result in Table 5 can clearly reject the null hypothesis in favor of taste-based discrimination, we tentatively conclude that status-based screening at the IETF is functional. In other words, IETF participants are not applying different publication standards to different types of author. Still, it will be interesting to revisit the academic journal and patent results in a few years — when there has been enough time for more cites to accumulate.

The other broad conclusion we draw from Table 5 is that one should be very cautious when interpreting these citations-based outcome tests. Perhaps the most striking result in the table is the large *positive* correlation between status and RFC citations, which disappears when we examine citations from outside the IETF community. If RFC citations reflect taste-based discrimination in the citing process, the outcome tests would be biased against finding taste-based discrimination in the screening stage. Nevertheless, outcome tests have been overlooked in the recent literature on status, and may provide a way to test alternative mechanims or explanations for status effects in settings where there is a particularly clean measure of unbiased *ex post* performance.

6 Conclusion

Many authors have written about the importance of labels and identity. Perhaps the most famous statement of the hypothesis that a name does not (or should not) matter belongs to Shakespeare:

> What's in a name? that which we call a rose
> By any other name would smell as sweet;

So Romeo would, were he not Romeo call'd,
Retain that dear perfection which he owes
Without that title.

This paper presents evidence that Juliet was wrong, at least within the context of Internet standards development. We exploit a unique natural experiment created by the fact that *et al* obscures the names of some authors who nevertheless contribute to proposals brought before the Internet Engineering Task Force. We find that when *et al* obscures the name of a high-status author — specifically a current or former IETF Working Group chair — there is a significant drop in the publication rate relative to the case where *et al* obscures the name of a low-status author.

These results provide statistical evidence of Merton's Matthew Effect in the novel organizational context of open standards development. Our estimates suggest that the impact of name-based signals is quite large in this setting, explaining up to two-thirds of the benefits of having a high-status author on the proposal. This is especially surprising given the "weak" nature of the treatment condition: even when *et al* obscures an author's name on the email announcement, it is relatively easy to find them by downloading the relevant proposal.

We ask what gives rise to this Matthew Effect; specifically, do author names serve as a signal of quality, or is there simply a "taste" for status that leads to the application of a double-standard that favors high-status authors? Our findings suggest that names are primarily a signal of expected quality: the screening process is one of statistical rather than taste-based discrimination. This conclusion is based on three pieces of evidence. First, there are no status-signaling effects in the sample of Internet Drafts submitted to Working Groups. Second, publication rates respond to other information that gets obscured by *et al* (i.e. how many RFCs an author has published) and not just status cues. And thirdly, citations to published RFCs suggest relatively little difference in the average quality of work produced by high and low status authors.

Our paper also explores the role of attention at early stages of the IETF publication process. In particular, we show that proposals from high-status authors generate more conversation among IETF participants, but only when their name appears on the announcement message. This suggests that increasing returns to attention may be a mechanism that explains how our relatively weak treatment condition leads to a substantial change in publication rates. More generally, it suggests a source of positive feedback from an actor's initial status or position to the underlying quality of their work. This type of feedback loop reconciles some of the tension between sociological theories that emphasize increasing returns to status, and economic models of signaling or reputation, where signals are only used when they convey information about the underlying quality of the sender.

Finally, our analysis highlights the enduring importance of Merton's proposition that, "it is important to consider the social mechanisms that curb or facilitate the incorporation of would-be contributions into the domain of science." The last forty years have seen a remarkable decline in the costs of accessing knowledge, often based on technologies developed within the IETF. At the same time, there has been a tremendous growth in the volume of new ideas. Our analysis suggests that status and social networks remain key to garnering attention for new ideas in such an environment.

References

Ai, C. & Norton, E. C. (2003). Interaction Terms in Logit and Probit Models. *Economics Letters*, *80*(1), 123–129.

Ayres, I. (2002). Outcome Tests of Racial Disparities in Police Practices. *Justice Research and Policy*, *4*, 132–142.

Ayres, I. & Waldfogel, J. (1994). A Market Test for Race Discrimination in Bail Setting. *Stanford Law Review*, *46*(5), 987–1047.

Becker, G. S. (1993). Nobel Lecture: The Economic Way of Looking at Behavior. *The Journal of Political Economy*, *101*(3), 385–409.

Benjamin, B. A. & Podolny, J. M. (1999). Status, Quality, and Social Order in the California Wine Industry. *Administrative Science Quarterly*, *44*(3), 563–589.

Berger, J., Rosenholtz, S., & Zelditch, M. (1980). Status Organizing Processes. *Annual Review Of Sociology*, *6*, 479–508.

Blank, R. M. (1991). The Effects of Double-Blind versus Single-Blind Reviewing: Experimental Evidence from The American Economic Review. *The American Economic Review*, *81*(5), 1041–1067.

Bothner, M. S. & Haynes, R. (2009). When do Matthew Effects Occur? Technical report, University of Chicago Working Paper.

Bothner, M. S., Kim, Y.-K., & Smith, E. B. (2009). How Does Status Affect Performance? Status as an Asset Versus Status as a Liability in the PGA and NASCAR. Working paper, University of Chicago Working Paper.

Burris, V. (2004). The Academic Caste System: Prestige Hierarchies in PhD Exchange Networks. *American Sociological Review*, *69*(2), 239–264.

Burt, R. S. (2004). Structural Holes and Good Ideas. *American Journal of Sociology*, *110*, 349–399.

DellaVigna, S. (2008). Psychology and Economics: Evidence from The Field. *Journal of Economic Literature*, *47*(2), 315–372.

Esteves-Sorenson, C. (2009). Micro-Costs: Inertia in Television Viewing. Technical report, Yale School of Management Working Paper.

Feldman, M. S. & March, J. G. (1981). Information in Organizations as Signal and Symbol. *Administrative Science Quarterly*, *26*(2), 171–186.

Fleming, L. & Waguespack, D. (2009). Scanning the Commons? Evidence on the benefits to startups of participation in open standards development. *Management Science*, *55*(2), 210–223.

Fuller, B. (2006). Phone Interview - November 14 2006.

Goldin, C. & Rouse, C. (2000). Orchestrating Impartiality: The Impact of “Blind” Auditions on Female Musicians. *The American Economic Review*, *90*(4), 715–741.

Gould, R. (2002). The origins of status hierarchies: A formal theory and empirical test. *American Journal of Sociology*, *107*(5), 1143–1178.

Hall, B., Jaffe, A., & Trajtenberg, M. (2005). Market value and patent citations. *Rand Journal Of Economics*, *36*(1), 16–38.

Knowles, J., Persico, N., & Todd, P. (2001). Racial Bias in Motor Vehicle Searches: Theory and Evidence. *The Journal of Political Economy*, *109*(1), 203–229.

Kreps, D. M. & Wilson, R. (1982). Reputation and Imperfect Information. *Journal of Economic Theory*, *27*, 253–279.

Merton, R. K. (1968). The Matthew Effect in Science. *Science*, *159*, 56–63.

Morris, S. & Shin, H. S. (2002). Social Value of Public Information. *The American Economic Review*, *92*(5), 1521–1534.

Mouw, T. (2006). Estimating the causal effect of social capital: A review of recent research. *Annual Review of Sociology*, *32*, 79–102.

Mowery, D. & Simcoe, T. (2002). Is the Internet a US invention? An economic and technological history of computer networking. *Research Policy*, *31*(8-9), 1369–1387.

Phillips, D. & Zuckerman, E. (2001). Middle-status conformity: Theoretical restatement and empirical demonstration in two markets. *American Journal of Sociology*, *107*(2), 379–429.

Podolny, J. (1994). Market Uncertainty And The Social Character Of Economic Exchange. *Administrative Science Quarterly*, *39*(3), 458–483.

Podolny, J. & Stuart, T. (1995). A Role-Based Ecology Of Technological-Change. *American Journal Of Sociology*, *100*(5), 1224–1260.

Podolny, J. M. (1993). A Status Based Model of Market Competition. *American Journal of Sociology*, *98*(4), 829–872.

Podolny, J. M. (2005). *Status Signals: A Sociological Study of Market Competition.* Princeton University Press.

Podolny, J. M. & Phillips, D. J. (1996). The Dynamics of Organizational Status. *Industrial and Corporate Change*, *5*, 453–472.

Powell, W., Koput, K., & SmithDoerr, L. (1996). Interorganizational collaboration and the locus of innovation: Networks of learning in biotechnology. *Administrative Science Quarterly*, *41*(1), 116–145.

Rao, H., Greve, H., & Davis, G. (2001). Fool's gold: Social proof in the initiation and abandonment of coverage by Wall Street analysts. *Administrative Science Quarterly*, *46*(3), 502–526.

Ridgeway, C. & Erickson, K. (2000). Creating and spreading status beliefs. *American Journal Of Sociology*, *106*(3), 579–615.

Simcoe, T. (2008). Standard Setting Committees. *Working Paper.*, *(http://ssrn.com/abstract=899595).*

Sorenson, O. & Waguespack, D. M. (2006). Social structure and exchange: Self-confirming dynamics in Hollywood. *Administrative Science Quarterly*, *51*(4), 560–589.

Spence, M. (1973). Job Market Signaling. *Quarterly Journal of Economics*, *87*(3), 355–374.

Stewart, D. (2005). Social status in an open-source community. *American Sociological Review*, *70*(5), 823–842.

Stuart, T., Hoang, H., & Hybels, R. (1999). Interorganizational endorsements and the performance of entrepreneurial ventures. *Administrative Science Quarterly*, *44*(2), 315–349.

Washington, M. & Zajac, E. (2005). Status evolution and competition: Theory and evidence. *Academy of Management Journal*, *48*(2), 282–296.

Tables and Figures

Table 1: Variable Definitions

Variable Name	Definition
Published as RFC	Indicator: Internet Draft (ID) is published as an RFC
Email Lists	Number of email lists where ID is mentioned at least once
Replies	Number of reply messages (from any listserv) that mention ID
Revisions	Number of times ID is submitted to IETF
RFC Cites	Citations from future RFCs to focal RFC
Patent Cites	Citations from US patents to focal RFC
Article Cites	Citations from ISI academic journal articles to focal RFC
Et Al Dummy	Indicator: One or more authors not listed in ietf-announce
WG Chair Author	Indicator: At least one author is a past or current WG chair
Unlisted WG Chair	Indicator: WG Chair Author = 1 & No Chair listed in ietf-announce
Lead Chair	Indicator: First ID author is a past or current WG chair
Published RFCs	Log one plus sum of RFCs previously published by all ID authors
Unlisted RFCs	Log one plus sum of RFCs published by authors not in ietf-announce
Authors	Total number of authors listed on draft
Log pages	Log of number of pages in Internet Draft
Published Author	Indicator: At least one author has published an RFC
Future RFCs	Log one plus sum of RFCs published in future by all draft authors
Intl Author	Indicator: At least one international author
Multi-sponsor	Indicator: Draft authors affiliated with more than one organization
Days-to-meeting	Log count of days until next IETF meeting
Draft Year	Year when draft is first submitted to IETF
Publication Year	Year when draft is Published as RFC

Table 2: Summary Statistics

	Individual IDs			Working Group IDs		
Variable	Mean	S.D.	Obs.	Mean	S.D.	Obs.
Published as RFC	0.069	0.254	3719	0.468	0.499	1326
Email Lists	1.409	1.748	3719	5.096	5.573	1326
Replies	4.782	9.370	2534	12.237	21.641	1192
Revisions	1.883	1.766	3719	4.505	3.562	1326
RFC Cites	4.352	9.404	250	7.390	13.203	602
Patent Cites	0.368	1.296	250	0.816	3.153	602
Article Cites	0.368	1.409	250	0.982	3.686	602
Et Al Dummy	0.195	0.396	3719	0.137	0.344	1326
WG Chair Author	0.259	0.438	3719	0.500	0.500	1326
Unlisted WG Chair	0.034	0.182	3719	0.041	0.199	1326
Lead Chair	0.175	0.380	3719	0.312	0.464	1326
Published RFCs	0.803	1.110	3719	1.563	1.196	1326
Unlisted RFCs	0.109	0.481	3719	0.157	0.559	1326
Authors	2.101	1.202	3719	2.347	1.288	1326
Log pages	2.396	0.951	3719	2.696	1.059	1326
Published Author	0.436	0.496	3719	0.771	0.421	1326
Future RFCs	0.683	0.882	3719	1.326	0.965	1326
Intl Author	0.401	0.490	3719	0.325	0.469	1326
Multi-sponsor	0.354	0.478	3719	0.563	0.496	1326
Days-to-meeting	3.654	0.684	3719	3.753	0.716	1326
Publication Year	2001.6	1.079	3719	2001.5	1.096	1326

Table 3: Identity as a Signal[†]

Linear Probability Models of ID Publication
Unit of Observation = Internet Draft
Dependent Variable = Published as RFC

Sample	**All Individual IDs**			**All WG IDs**	
WG Chair Author	0.101**		0.030*	0.179**	0.010
	(0.01)		(0.02)	(0.03)	(0.03)
Unlisted WG Chair	-0.066**		-0.009	0.001	0.039
	(0.02)		(0.02)	(0.08)	(0.08)
Published RFCs		0.051**	0.040**		0.094**
		(0.01)	(0.01)		(0.01)
Unlisted RFCs		-0.033**	-0.031**		-0.039
		(0.01)	(0.01)		(0.04)
Et Al Dummy	-0.051**	-0.027*	-0.016	-0.100*	-0.111*
	(0.01)	(0.01)	(0.01)	(0.04)	(0.05)
Log Pages			0.013**		0.053**
			(0.00)		(0.01)
Intl Author			-0.019*		0.009
			(0.01)		(0.03)
Multi-sponsor			0.007		0.039
			(0.01)		(0.05)
Days-to-meeting			0.047**		0.133**
			(0.01)		(0.02)
Constant	0.055**	0.084**	-0.162**	0.392**	-0.271**
	(0.00)	(0.03)	(0.05)	(0.02)	(0.09)
Author Count Effects	N	Y	Y	N	Y
IETF Meeting Effects	N	Y	Y	N	Y
Observations	3719	3719	3719	1326	1326
R-squared	0.036	0.063	0.082	0.035	0.171
Mean of DV	0.069	0.069	0.069	0.468	0.468

Robust standard errors in parentheses; +10% significance; *5% significance; **1% significance. [†]See Table A1 for additional models that control for rank of Unlisted Chair in list of authors.

Table 4: Identity and Attention†

OLS Regressions

Unit of Observation = Internet Draft

Sample = Individual Drafts

Dependent Variable	Email Lists		Replies		Revisions	
WG Chair Author	0.678**	0.355**	2.294**	1.502*	0.671**	0.342**
	(0.09)	(0.09)	(0.51)	(0.60)	(0.09)	(0.10)
Unlisted WG Chair	-0.480**	-0.447**	-1.116	-0.762	-0.603**	-0.631**
	(0.17)	(0.17)	(1.10)	(1.07)	(0.15)	(0.16)
Et Al Dummy	0.108	-0.442**	-0.322	-1.365*	-0.209**	-0.702**
	(0.07)	(0.11)	(0.47)	(0.62)	(0.07)	(0.11)
Published Author		0.368**		0.794+		0.350**
		(0.07)		(0.46)		(0.06)
Log Pages		0.234**		0.312+		0.291**
		(0.03)		(0.19)		(0.03)
Intl Author		-0.027		0.162		-0.107+
		(0.06)		(0.41)		(0.06)
Multi-sponsor		0.259**		1.313**		0.354**
		(0.08)		(0.49)		(0.07)
Days-to-meeting		-0.174**		-0.042		0.148**
		(0.04)		(0.29)		(0.05)
Constant	1.229**	1.741**	4.236**	2.347+	1.770**	0.848*
	(0.03)	(0.34)	(0.23)	(1.30)	(0.03)	(0.35)
Author Count Effects	N	Y	N	Y	N	Y
IETF Meeting Effects	N	Y	N	Y	N	Y
Observations	3719	3719	2534	2534	3719	3719
R-squared	0.026	0.089	0.012	0.040	0.029	0.101
Mean of D.V.	1.409	1.409	4.782	4.782	1.883	1.883

Robust standard errors in parentheses; +10% significance; *5% significance; **1% significance.†See TableA2 for similar results based on Published RFCs as a proxy for status (reputation).

Table 5: Status and Citations

Poisson Regressions

Unit of Observation = RFC

Sample = Individual IDs (1995-2003) Published as RFC

Dependent Variable	RFC Cites		Patent Cites		Article Cites	
WG Chair Author	1.995**	0.654	0.223	0.310	-0.192	-0.732*
	(0.68)	(0.44)	(0.31)	(0.36)	(0.28)	(0.30)
Published RFCs		0.344**		0.042		0.242+
		(0.12)		(0.12)		(0.14)
Future RFCs		1.163**				
		(0.44)				
Log Pages		-0.237		0.498**		0.140
		(0.37)		(0.13)		(0.17)
Intl Author		0.821*		-0.458		-0.854**
		(0.39)		(0.28)		(0.32)
Multi-sponsor		0.831*		-0.288		0.531
		(0.36)		(0.43)		(0.36)
Constant	0.415	-0.303	-3.631**	-4.969**	-2.374**	-2.741**
	(0.57)	(0.64)	(1.02)	(1.13)	(0.62)	(0.99)
Author Count Effects	Y	Y	Y	Y	Y	Y
Publication Year Effects	Y	Y	Y	Y	Y	Y
Observations	479	479	479	479	479	479
Mean of DV	10.89	10.89	1.34	1.34	0.84	0.84

Robust standard errors in parentheses; $^{+}$10% significance; *5% significance; **1% significance.

Figure 1: The IETF Publication Process

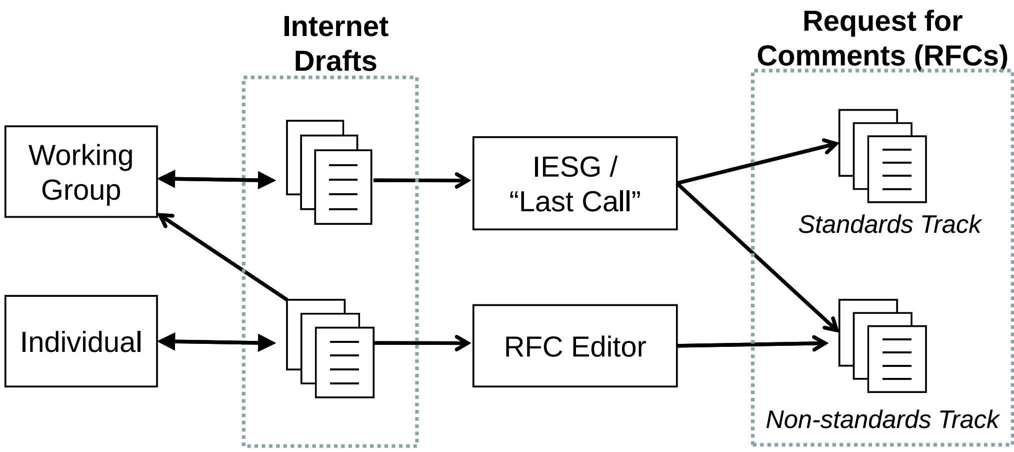


Figure 2: Author Names in ietf-announce and Internet Drafts

Top panel shows a typical ietf-announce message listing all author names. Bottom panel shows front page and authors' addresses as they appear in the Internet Draft.

A New Internet-Draft is available from the on-line Internet-Drafts directories.

Title : Basic Network Mobility Support
Author(s) : R. Wakikawa, K. Uehara, K. Mitsuya, T. Ernst
Filename : draft-wakikawa-nemo-basic-00.txt
Pages : 21
Date : 2003-2-18

This draft proposes a solution for Basic Network Support. It proposes Mobile IPv6 extensions as advocated by the NEMO working group. Our solution differs from Prefix Scope Binding Update …

INTERNET DRAFT
18 Feb 2003

Ryuji Wakikawa
Keisuke Uehara
Koshiro Mitsuya
Thierry Ernst
Keio University and WIDE

Basic Network Mobility Support
draft-wakikawa-nemo-basic-00.txt

… (BODY TEXT) …

Authors' Addresses

Ryuji Wakikawa
Keio University and WIDE
5322 Endo Fujisawa Kanagawa
252-8520 JAPAN
Phone: +81-466-49-1394
Fax: +81-466-49-1395
EMail: ryuji@sfc.wide.ad.jp

Koshiro Mitsuya
Keio University and WIDE
5322 Endo Fujisawa Kanagawa
252-8520 JAPAN
Phone: +81-466-49-1394
Fax: +81-466-49-1395
EMail: mitsuya@sfc.wide.ad.jp

Keisuke Uehara
KEIO University and WIDE
5322 Endo Fujisawa Kanagawa
252-8520 JAPAN
Phone: +81-466-49-1394
Fax: +81-466-49-1395
EMail: kei@wide.ad.jp

Thierry Ernst
Keio University and WIDE
5322 Endo Fujisawa Kanagawa
252-8520 JAPAN
Phone: +81-466-49-1394
Fax: +81-466-49-1395
EMail: ernst@sfc.wide.ad.jp

Figure 3: *Et Al* and Missing Author Names

Example on left shows an ietf-announce message where *et al* obscures three author names: W. Hardaker, D. Harrington and M. Stiemerling. Example on right shows an ietf-announce message that omits two author names: J. Hand and G. Swallow.

A New Internet-Draft is available from the on-line Internet-Drafts directories.

Title : Middlebox Communications (MIDCOM) Protocol Managed Objects
Author(s) : M. Barnes et al.
Filename : draft-barnes-midcom-mib-01.txt
Pages : 16
Date : 2003-7-1

This document describes and defines the managed objects for dynamic configuration of middleboxes. The scope of the middleboxes to which these managed objects apply is limited to NATs and Firewalls …

Internet Draft — M. Barnes
Document: draft-barnes-midcom-mib-01.txt — Nortel Networks
Wes Hardaker
Sparta
D. Harrington
Enterasys Networks
M. Stiemerling
Category: Standards Track — NEC Europe Ltd.
Expires: December 2003 — June 2003

Middlebox Communications (MIDCOM) Protocol Managed Objects

… (BODY TEXT) …

Authors' Address

Mary Barnes
Nortel Networks
2380 Performance Drive
Richardson, TX 75082
USA
Phone: 1-972-684-5432
Email: mbarnes@nortelnetworks.com

…..

David Harrington, Co-chair SNMPv3 WG
Enterasys Networks

….

A New Internet-Draft is available from the on-line Internet-Drafts directories.

Title : End-to-End VoIP over MPLS Header Compression
Author(s) : J. Ash, B. Goode
Filename : draft-ash-e2e-vompls-hdr-compress-01.txt
Pages : 0
Date : 2003-3-6

VoIP over MPLS typically uses the encapsulation voice/RTP/UDP/IP/MPLS. For an MPLS VPN, the packet header is at least 48 bytes, while the voice payload is typically no more than 30 bytes. VoIP over MPLS header…

Network Working Group — Jerry Ash
Internet Draft — Bur Goode
<draft-ash-e2e-vompls-hdr-compress-01.txt> — Jim Hand
Expiration Date: October 2003 — AT&T
George Swallow
Cisco Systems, Inc.

March, 2003

End-to-End VoIP over MPLS Header Compression

… (BODY TEXT) …

7. Authors' Addresses

Jerry Ash
AT&T
Room MT D5-2A01
200 Laurel Avenue
Middletown, NJ 07748, USA
Phone: +1 732-420-4578
Email: gash@att.com

….

George Swallow
Cisco Systems, Inc.
250 Apollo Drive Chelmsford, MA 01824
Phone: +1 978 497 8143
Email: swallow@cisco.com
…

Figure 4: New Internet Draft Submissions

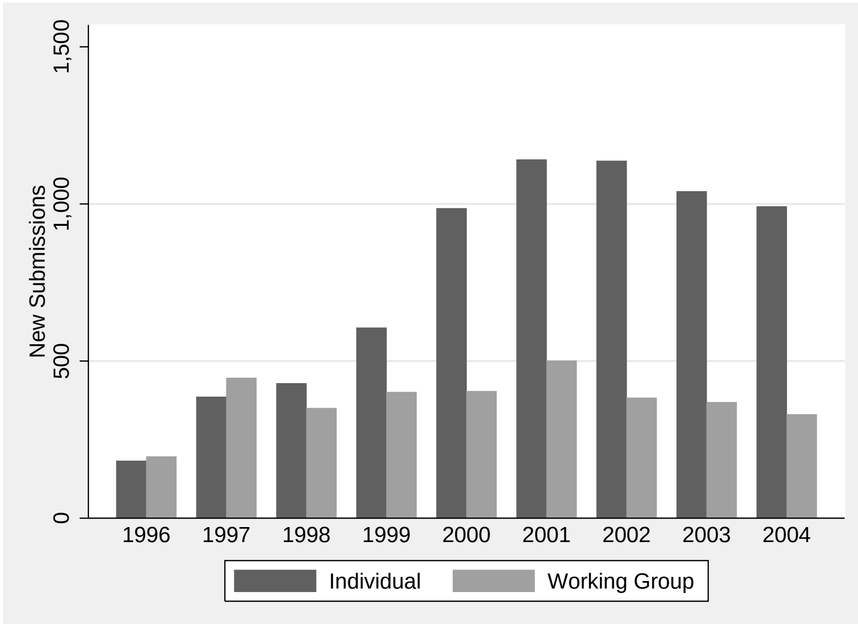


Figure 5: Probability of Et Al

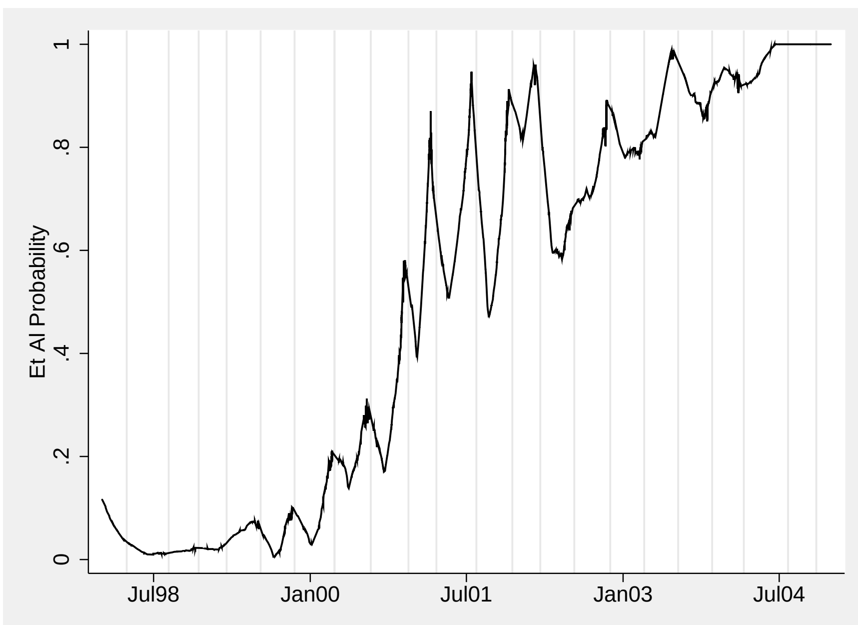
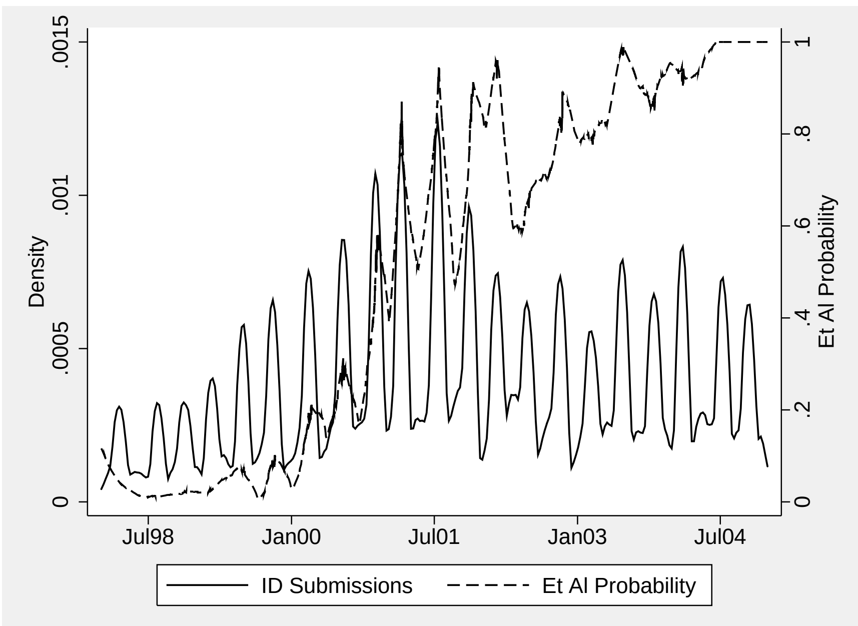


Figure 5 is based on a lowess regression where sample includes all IDs with 2 or more authors. Vertical lines indicate IETF Meeting dates.

Figure 6: Et Al and ID Submission Rates



Appendix: Robustness Checks

Table A1: Status and Reputation

Logistic Regressions
(Marginal Effects)

Unit of Observation = Internet Draft

Dependent Variable = WG Chair Author

Sample	All Drafts	Single Author
Published RFCs	0.244**	0.236**
	(0.01)	(0.02)
Two Author	-0.052**	
	(0.02)	
Three Author	-0.041*	
	(0.02)	
Four Author	0.013	
	(0.03)	
Five Author	0.006	
	(0.04)	
Year Effects	Y	Y
Observations	3719	1554
Mean of DV	0.19	0.16

Robust standard errors in parentheses; $^{+}$10% significance; *5% significance; **1% significance.

Table A2: Robustness of Signaling Effect

Linear Probability Models of Individual ID Publication

Unit of Observation = Internet Draft

Dependent Variable = Published as RFC

Sub-Sample	All Drafts	2-5 Authors & No Lead Chair	2 Authors & No Lead Chair	3 Authors & No Chair 1-2	All Drafts	WG Chair Author & No Lead Chair
WG Chair Author	0.070**	0.075**	0.067+	0.067	0.070**	
	(0.01)	(0.03)	(0.04)	(0.07)	(0.01)	
Unlisted Chair2†	-0.105**	-0.086*	-0.062			
	(0.03)	(0.03)	(0.04)			
Unlisted Chair3	-0.065	-0.062		-0.070		
	(0.04)	(0.05)		(0.08)		
Unlisted Chair4	-0.051	-0.060				
	(0.04)	(0.04)				
Unlisted Chair5	-0.063*	-0.080*				
	(0.03)	(0.04)				
Unlisted WG Chair					-0.058*	-0.118**
					(0.03)	(0.04)
Et Al Dummy	-0.028*	-0.039**	-0.036**	-0.060**	-0.025+	
	(0.01)	(0.01)	(0.01)	(0.02)	(0.01)	
Missing Et Al††					-0.012	
					(0.02)	
Missing * WG Chair					-0.010	
					(0.04)	
Published Author	0.042**	0.029*	0.035+	0.044+	0.042**	-0.001
	(0.01)	(0.01)	(0.02)	(0.02)	(0.01)	(0.05)
Log pages	0.014**	0.016**	0.019*	0.012	0.014**	0.032*
	(0.00)	(0.01)	(0.01)	(0.01)	(0.00)	(0.02)
Intl Author	-0.021*	-0.021*	-0.020	-0.024	-0.021*	-0.068*
	(0.01)	(0.01)	(0.02)	(0.02)	(0.01)	(0.03)
Multi-sponsor	0.012	-0.000	-0.004	0.000	0.012	0.037
	(0.01)	(0.01)	(0.02)	(0.02)	(0.01)	(0.05)
Days-to-meeting	0.049**	0.028**	0.037**	0.026+	0.049**	0.096**
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.03)
Constant	-0.164**	-0.079	-0.078	-0.049	-0.161**	-0.211
	(0.05)	(0.06)	(0.06)	(0.06)	(0.05)	(0.14)
Author Count Effects	Y	Y	Y	Y	Y	Y
IETF Meeting Effects	Y	Y	Y	Y	Y	Y
Observations	3719	1887	835	507	3719	315
R-squared	0.072	0.061	0.076	0.112	0.072	0.144
χ^2 P-value*	0.00	0.13			0.69	

Robust standard errors in parentheses; +10% significance; *5% significance; **1% significance. †Unlisted ChairX = 1 iff the X^{th} author is a WG chair and is not listed in the ietf-announce message. ††Missing EtAl = 1 iff the ietf-announce message does not list all authors and does not use *et al*to acknowledge missing names. *Null hypothesis is that all Unlisted ChairX coefficients (or Missing *et al*and Missing * WG Chair) are jointly equal to zero.

Table A3: Attention and Past Publications

OLS Regressions Unit of Observation = Internet Draft Sample = Individual Drafts						
Dependent Variable	Email Lists		Replies		Revisions	
Published RFCs	0.307**	0.156**	1.025**	0.621+	0.382**	0.399**
	(0.04)	(0.06)	(0.21)	(0.35)	(0.04)	(0.06)
Unlisted RFCs	-0.161*	-0.172*	-0.821*	-0.856*	-0.133	-0.175
	(0.08)	(0.07)	(0.37)	(0.37)	(0.13)	(0.12)
Et Al Dummy	0.129+	-0.407**	-0.031	-0.952	-0.226**	-0.667**
	(0.07)	(0.11)	(0.50)	(0.62)	(0.06)	(0.10)
Published Author		0.265*		0.546		-0.198+
		(0.11)		(0.70)		(0.12)
Log Pages		0.229**		0.301		0.285**
		(0.03)		(0.19)		(0.03)
Intl Author		-0.050		0.049		-0.089
		(0.06)		(0.40)		(0.06)
Multi-sponsor		0.270**		1.415**		0.302**
		(0.08)		(0.49)		(0.07)
Days-to-meeting		-0.184**		-0.079		0.134**
		(0.04)		(0.29)		(0.04)
Constant	1.155**	1.823**	4.020**	2.635*	1.634**	0.915**
	(0.03)	(0.34)	(0.23)	(1.30)	(0.03)	(0.33)
Author Count Effects	N	Y	N	Y	N	Y
IETF Meeting Effects	N	Y	N	Y	N	Y
Observations	3719	3719	2534	2534	3719	3719
R-squared	0.034	0.088	0.013	0.039	0.058	0.113
Mean of D.V.	1.409	1.409	4.782	4.782	1.883	1.883

Robust standard errors in parentheses; $^{+}$10% significance; *5% significance; **1% significance.