

# Status, Quality and Attention: What's in a (Missing) Name? \*

Timothy S. Simcoe

*Boston University School of Management and NBER*

Dave M. Waguespack

*Robert H. Smith School of Business, University of Maryland*

May 5, 2010

---

\*We gratefully acknowledge financial support from the NET Institute ([www.netinst.com](http://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, Jóhanna K. Birnir, Scott Bradner, Brent Goldfarb, Ben Hallen, Olav Sorenson and David Tan for helpful comments. Address for correspondence: Boston University School of Management, 595 Commonwealth Avenue, Boston, MA 02215. E-mail: [tsimcoe@bu.edu](mailto:tsimcoe@bu.edu)

# Status, Quality and Attention: What's in a (Missing) Name?

## Abstract

How much are we influenced by an author's identity when evaluating their work? This paper addresses this question 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). By comparing the effect of obscuring high versus low status author names, we measure the impact of status signals on the IETF publication process. Our results suggest that name-based signals can explain up to three-quarters of the difference in publication outcomes across status cohorts. However, this signaling effect disappears for a set of pre-screened proposals that receive more scrutiny than a typical submission. We also show that working papers from high status authors receive more attention on electronic discussion boards, which may help in developing these ideas and bringing them forward to publication. JEL Codes: L1, O3

# 1 Introduction

Sociologists studying markets have long emphasized that identity, particularly the visible position within a status hierarchy, matters independently of underlying quality for competitive outcomes. In transactions where there is uncertainty about the underlying quality of the focal producer or their products, consumer decisions about whom to transact with and how much the product is worth are influenced by social affiliation cues (Podolny, 1993, 1994). Prominent individuals and organizations with superior *status signals* benefit from these acts of deference (Merton, 1968; Podolny, 1993). Producer identity consequently matters, the argument goes, for a wide variety of decisions, from whose paper to publish to which job applicant to hire, or how much to pay for a particular bottle of wine.

Advancing this line of research, however, is severely hindered by the dynamic relationship between status, quality, and performance in real world settings. For instance, Podolny & Phillips (1996) argue that status is partially determined by innate ability, and that high status producers benefit both from greater perceived quality and from greater ability to produce high quality goods due to the superior resources they receive. The reciprocal link between status and quality consequently makes it hard to dis-entangle their effects. Multivariate regression can produce misleading results — even if one includes a *perfect* control for quality — because the two constructs are jointly determined. The primary contribution of this paper is in detailing a testing approach that partly circumvents this problem by exploiting random variance in status signals, as opposed to observed variation in status itself.

Our basic approach will be familiar to anyone who has participated in the peer review process; and is nicely illustrated by the following anecdote about a paper submitted to the British Association for the Advancement of Science by the third Lord Rayleigh, who went on to win the 1904 Nobel Prize in Physics:

“His [Rayleigh’s] name was either omitted or accidentally detached, 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.” (Strutt, 1968, p.228 as quoted in Merton (1968)).

In the Rayleigh case, we have solid evidence that identity effects are not simply due to unobserved quality, because removing Rayleigh’s name had no bearing on the innate merits of the manuscript. Barring such an accident, it is difficult to separate the status signal from actual quality in observational data, particularly as we expect that high quality producers will attempt to acquire (and signal) greater status. The most talented scientists, for instance, may disproportionately seek education or employment at high status institutions, leading to a strong

correlation between status and latent quality. The Lord Rayleigh example also highlights the link between status and the allocation of a scarce resource, namely the critical attention of other scientists. If attention is critical for developing and improving new ideas, prominent scientists may have a cumulative advantage derived from superior resources, and not just from successive ratcheting up of perceived quality.

This study is essentially a large scale replication of the “Lord Rayleigh experiment” that emphasizes both identity-based signaling, and the increased attention allocated to high-status actors. Our data come from the Internet Engineering Task Force (IETF), a community of engineers and computer scientists who develop protocols used to run the Internet. All new protocols go through a non-blind peer review process, which begins with the IETF’s clerical staff posting a newly submitted manuscript to a web server, and describing it in a short broadcast email announcement. The connection to the Lord Rayleigh experiment is that between 2000 and 2003, the broadcast emails announcing new manuscripts often used the generic label *et al* instead of providing a complete list of author names. Both interview evidence and statistical tests indicate that the use of *et al* was unrelated to the quality of the paper or the social standing of the authors. We therefore treat name-removal via the use of *et al* as a natural experiment that allows us to measure the impact of identity-based signals independently of the latent quality of a manuscript or its authors.

We begin by examining the relationship between author names and publication decisions for 5413 papers submitted between 2000 and 2003. We report two main findings. First, concealing high status author names makes publication substantially less likely, but only among a pool of “general interest” manuscripts where submission rates are very high and the average rate of publication is quite low. In other words, identity matters independently of product attributes when search costs are high. By contrast, within a pool of manuscripts “pre-screened” as relevant to one of the IETF’s technical working groups the author name signal does not matter. Second, and just as importantly, we show that the conventional approach to estimating status effects is highly sensitive to model specification, and can produce large statistically significant effects in settings where our natural experiment fails to reject the null hypothesis that audiences are not responding to status signals.

We next examine the attention focused on authors prior to the publication decision for the general interest manuscripts. 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 electronic discussion boards, 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 critical attention to a new idea, which is useful for developing the concept and bringing it forward to publication (Podolny, 2005, p.26).

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, 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 affect the focal actor’s performance (Podolny & Phillips, 1996). Thus, Berger et al. (1980, p.479) define 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 affects 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).<sup>1</sup>

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

---

<sup>1</sup>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).

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 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, some basic questions remain unresolved. For example, in a review of social capital research, Mouw (2006) argues 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 two 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 between producer status and observed performance. 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 clever and reflect careful thinking about the testing context. For instance, in estimating the effect of producer status on the price of a bottle of wine Benjamin & Podolny (1999) control for bottle quality using scores from blind taste tests. However, if status and quality are highly correlated, this approach will always be sensitive to *untestable* assumptions about measurement and functional form. 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. Moreover, if one takes the position that both status and quality are latent constructs, a linear regression will produce biased estimates even if one has a perfect measure of perceived quality, since perceptions of quality and status are jointly determined.

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 this study, we focus on author names, as in the Lord Rayleigh example. Names provide a rather coarse signal, 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, we are able to measure the causal impact of the identity-based signal.

## 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 improvement in the quality of the focal actor’s products. The source of these endogenous improvements could be behavioral: Merton (1968) notes that the validation received by a star scientist may give them the self confidence required to tackle important problems; Phillips & Zuckerman (2001) posit that middle status actors are constrained by legitimacy concerns from taking risky actions; and Benjamin & Podolny (1999) find that high status wine producers select higher quality grapes. 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 enhance perceived quality.

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 differences in the behaviors or resources of high status actors, particularly *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. Of course, the same concerns with unobserved quality, reverse causation and simultaneity that hinder empirical tests of the status-performance relationship also apply to these intermediate outcomes. In the following sections we describe an empirical strategy that relies on random variation in status signals to address these concerns.

## 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; Waguespack & Fleming, 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 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.<sup>2</sup> 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.<sup>3</sup> 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

---

<sup>2</sup>In our data, Working Group IDs have a 47 percent publication rate, compared to 7 percent for individual submissions.

<sup>3</sup>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.

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 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 double 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. IETF authors are aware of the identity of others choosing to support, ignore, or criticize their work. 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 the 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.<sup>4</sup>

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 aca-

---

<sup>4</sup>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.

demic publishing is usually characterized by a clear hierarchy among journals within a field, all contributions to the IETF are published as RFCs.<sup>5</sup>

### 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 within 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).<sup>6</sup>

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

---

<sup>5</sup>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.

<sup>6</sup>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.

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.

The key assumption behind our empirical tests is that the IETF Secretariat’s use of *et al* is uncorrelated with the identity of individual authors. Importantly, this assumption is *testable*. In particular, we can examine the correlation between *et al* and all author-level explanatory variables in our analysis. If none of these control variables has a statistically significant correlation with *et al*, we should feel more confident in the maintained assumption that *unobserved* author characteristics (which, by definition, appear in the regression error-term) are also uncorrelated with the “treatment” condition.

To conduct this test of our central empirical assumption, we estimate a series of logit regressions for IDs with two or more authors (see Table A1). In each of these models, proximity to an IETF meeting and the number of authors on the ID were extremely significant. However, we found no evidence of a correlation between WG Chair authorship and *et al* usage, and in three out of four cases could not reject the joint null hypothesis that all observable ID characteristics had no impact on the probability of receiving an *et al*.

## 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](http://www.watersprings.org), and from the archived [ietf-announce](mailto:ietf-announce) mailing list. The estimation sample includes all IDs submitted between 2000, when the *et al* experiment began, and 2003.<sup>7</sup> We also exclude a small number of proposals with more than five listed authors, though including them does not alter our main results.<sup>8</sup>

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 4,013 individual IDs and 1,405 Working Group IDs in our estimation sample.

---

<sup>7</sup>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.

<sup>8</sup>Ninety-three percent of individual submissions and ninety percent of Working Group IDs had five authors or fewer.

## 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 7.2 percent for individual IDs and 49.7 percent for Working Group IDs. We also have several intermediate outcomes. *Replies* counts the number of “reply” messages that mention a focal ID within fourteen days of its initial submission.<sup>9</sup> *Email Lists* is a count of the email lists where a focal ID is discussed. Both variables were constructed by an automated 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. Based on these sample means, Working Group IDs receive more attention than individual submissions — about ten percent more replies, twice as many email lists and two more revisions .

The remaining rows 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.<sup>10</sup> Twenty-one percent of individual submissions and 23 percent of WG drafts receive an *et al*.

Our 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* (3.5 percent) is smaller than the product of *Et Al Dummy* (20.7 percent) and *WG Chair Author* (25.1 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 140 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

---

<sup>9</sup> “Reply” message counts provide a better measure of attention than total message counts because it excludes a variety of administrative emails that may mention an ID.

<sup>10</sup> In robustness tests, we find little difference between omitting names, and obscuring them with *et al*.

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.<sup>11</sup> Studies that attempt to measure status effects directly may require more refined measures, such as those based on positioning within a social network. Since we are measuring a *signaling* effect, and not the impact of a change in status, a single crude but salient proxy for individual reputation is arguably more suitable for our purposes.

Our main control for author experience or reputation is the variable *Published RFCs*, which is a logged count of the number of RFCs previously published by all authors who contribute to the ID. While there is no mechanical link between publishing RFCs and becoming a WG chair, Fleming & Waguespack (2007) 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 A2). This extremely strong 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 additional control variables. An average ID in our estimation sample has between two and three authors (*Authors*), and is roughly 11 pages long (*Log Pages*). 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. *Standards-track* and *Future RFCs* are used as control variables in the citation analysis.

## 4.2 Methods

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 simple approach would over-estimate the signaling effect if other factors produce a

---

<sup>11</sup>Though WG chairs cannot submit drafts to their own group, the political nature of their position highlights the role of name-removal in our analysis. While the variable *WG Chair Author* will capture any effect of high status individuals exercising political power (as well as unobserved quality, or other differences between chairs and non-chairs), there is little reason to think that political influence is hindered by receiving an *et al*: if anything, we expect that chicanery is aided by secrecy.

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  $\lambda_t$  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 ( $\beta_1$ ) will not equal the cross-partial derivative of the expected probability of publication (Ai & Norton, 2003).

In equation (1),  $\beta_1$  measures the signaling effect. The change in publication probability associated with having a high-status author ( $\beta_2$ ) may reflect a variety of factors, such as latent proposal quality or author-experience. Ideally, we would let  $\beta_1$  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. For instance, 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.

## 5 Results

Tables 3 and 4 present our main results. Table 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.

### 5.1 Identity as a Signal

Table 3 shows that name-based signaling has a significant impact on individual ID publication rates, but no measurable 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*.<sup>12</sup> The results show that individual IDs are 8.4 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 by 6.6 percentage points when the author’s name is obscured or omitted. Thus, name-based signaling accounts for roughly 78 percent of the total benefits of WG chair authorship. Both the main effect and the signaling effect are large in comparison to the 5.8 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 2.6 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 adds a variety of controls to our baseline specification, including a complete set of meeting-date and author-count fixed effects. If the *et al* process is truly uncorrelated with observable features of the ID, this should have no impact on our estimates of the signaling effect. And indeed, the coefficient on *Unlisted WG Chair* remains negative,

---

<sup>12</sup>In this specification, there is no possibility that predicted values fall outside the unit interval, even for a linear probability model.



statistically significant, and within the 95 percent confidence interval for our previous estimate. The main effect of *WG Chair Author* drops by roughly fifty percent, due to the inclusion of the highly correlated variable *Published RFCs*. In unreported models, we find that controlling for the number of RFCs published by *unobserved* authors produces a similar effect: the coefficient on *Unlisted WG Chair* falls sharply, but we can strongly reject the joint null hypothesis that both unobserved chairs and unobserved RFCs have no impact on publication rates. The additional control variables in column 2 are jointly significant, and the model R-squared is more than twice that of our first model. The main effect of *Et Al Dummy* drops to zero once *Days-to-meeting* is introduced to control for any congestion effects. We find that having an international author is associated with a small decline in publication rates. However, neither proposal size (*Log Pages*) nor author diversity (*Multi-Sponsor*) has a measurable impact on outcomes.

Since the first two columns of Table 3 report our main finding, we subject them to a variety of robustness checks (see Appendix Table A3). First, we show that similar results can be obtained as the marginal effects of a logit specification. We then show that it makes no difference whether a chair’s name is obscured by *et al* or simply omitted from the ietf-announce message. As a third robustness check, we allow the *Unlisted Chair* coefficient to vary according to a chair’s position in the list of authors. That model strongly rejects the hypothesis that all of the unlisted chair coefficients are zero ( $p=0.00$ ), which is reassuring, since one might worry about a correlation between *Unlisted Chair* and latent quality if authors are only obscured when they make a smaller contribution to the proposal. Next, we show that our main results are robust to dropping single authored drafts from the estimation sample. Finally, we show that the magnitude of the signaling effect is unchanged if we restrict the sample to IDs with two authors and no chair in the lead author position, though in this case our estimates lose statistical significance.

Returning to Table 3, the last two models 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 3 uses the very simple specification used in the first column. While we find a somewhat larger effect for *WG Chair Author*, the point estimate on the status-signaling parameter is half that of the individual IDs, and statistically indistinguishable from zero. This is consistent with the idea that status effect are more pronounced in environments characterized by high uncertainty. The results in column 3 also make an important methodological point. In particular, there is a large, positive and statistically significant coefficient on our proxy for status, even though we conclude that there is in fact no signaling. The results in column 4 show a similar pattern after introducing a full set of controls. Our point estimate of the status signaling effect drops to almost zero. As with individual IDs, the

main effect of *WG Chair Author* drops substantially when *Published RFCs* are introduced. Interestingly, we find that the coefficient on author diversity (*Multi-Sponsor*) becomes statistically significant in the Working group setting, where we might expect politics to play a larger role in the publication process.

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 three-quarters of the total benefit from 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 with high uncertainty. In particular, the level of uncertainty may not be an exogenous factor, but rather a function of how much attention individual IETF participants, and the entire community, pay to a particular set of 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. 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 *Replies*, *Email Lists* and *Revisions*. We interpret these measures as indicators of the amount of attention that an ID receives after it is released. Since we found no evidence of status signals for Working Group IDs, all of our results are based on the sample of individual submissions.<sup>13</sup>

The first two columns in Table 4 focus on *Replies* during the two week period after a proposal is first submitted. While the average individual ID is mentioned roughly 2.2 times, a proposal submitted by a WG chair author receives an additional 1.5 mentions. The main effect of *WG Chair Author* is highly significant. The coefficients on *Unlisted WG Chair* are large, accounting for between one-half and two-thirds of the main WG chair effect, but they are not precisely estimated. While the coefficient on *Unlisted WG Chair* is not statistically significant,

---

<sup>13</sup>Though we present OLS results for easy interpretation of the interaction coefficient, Appendix Table A4 shows that we obtain very similar results from a Poisson QML model (i.e. Poisson with robust standard errors to account for overdispersion).

the hypothesis tests in the bottom row of Table 4 do not reject the alternative null that IDs submitted by an unobserved WG chair receive no more than average attention.

The middle columns in Table 4 repeat this exercise using *Email Lists* as the outcome variable. We find that IDs from high-status authors are mentioned on 0.8 more email listservs than an average ID: a roughly 50 percent increase. In this model, the signaling effect — as measured by the coefficient on *Unlisted WG Chair* — is statistically significant at the ten percent level. Once again, we cannot reject the null hypothesis that proposals from an unobserved WG Chair receive no more attention than an average submission. The last two columns in Table 4 focus on *Revisions* and find a similar pattern. 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.

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.

## 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. Our method, which relies on manipulating the status signal as opposed to social position itself, shows that status effects are present for individual ID submissions, where there is high *ex ante* uncertainty and a scarcity of attention, but not for Working Group submissions, which typically receive a great deal more scrutiny. The standard approach of controlling for observable quality finds evidence of status effects in both cases.

These results provide statistical evidence of Merton’s Matthew Effect in the novel organizational context of open standards development. Our estimates suggest that under appropriate conditions, name-based signaling has a large impact on publication rates, explaining up to three-quarters 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 learn the identity of all authors by downloading the relevant proposal.

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 positive feedback loop between status, attention and quality 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. However, our paper does not address the important and contentious question of 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 privileged actors. While status is rewarded in either case, Merton (1968) calls the first mechanism “functional” since it ultimately selects for better outcomes. Our results suggest that name based screening at the IETF is functional because it only happens when search costs are high, and because we expect endogenous improvement in the quality of these manuscripts. The concept of status is also associated, however, with arguably non-

functional attributes such as power (Washington & Zajac, 2005) or celebrity (Rindova et al., 2005).

One approach for distinguishing between functional and non-functional screening involves decomposing measured status into distinct elements corresponding to perceived quality and social stratification (e.g. Piskorski, 2007; Jensen & Roy, 2008). Despite its prevalence in the literature, the decomposition approach is problematic in that it relies on strong assumptions about the validity of particular measures, and does not address the empirical problem of unobserved quality highlighted in this paper. An alternative approach utilizes the idea that functional and non-functional mechanisms have very different implications for long-run performance. In particular, Becker (1993) suggests that non-functional screening will lead to worse *ex post* performance within the favored group, since the application of a double-standard privileges the lowest quality actors within a favored group. This idea has been applied to distinguish between statistical (functional) and taste-based (non-functional) discrimination in a variety of settings (e.g. Ayres & Waldfogel, 1994; Knowles et al., 2001), and may provide future researchers with some traction on the difficult problem of distinguishing between functional and non-functional status mechanisms.

Regardless of whether identity-based signaling is functional or discriminatory, our results highlight 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.” in particular, we show that high-status actors to gather more attention for their ideas, and ultimately are more likely to get them published. This finding has significant implications for the dissemination of knowledge in an era of massive low cost distribution technology. Over the last forty years, a remarkable decline in the costs of accessing knowledge — often based on technologies developed within the IETF — has produced tremendous growth in the volume of new ideas. Our findings highlight the role of status and social networks remain in garnering attention for these ideas within a highly competitive intellectual environment, where attention is increasingly likely to constitute a scarce and valuable resource.

## References

- Ai, C. & Norton, E. C. (2003). Interaction Terms in Logit and Probit Models. *Economics Letters*, 80(1), 123–129.
- 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.
- 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. M. (2007). Brokerage, boundary spanning, and leadership in open innovation communities. *Organization Science*, 18(2), 165–180.
- 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.
- Jensen, M. & Roy, A. (2008). Staging exchange partner choices: When do status and reputation matter? *Academy of Management Journal*, 51(3), 495–516.

- 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.
- Piskorski, M. J. (2007). Positions of Power and Status: Reciprocity in the Venture Capital Industry. *Working Paper*.
- 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.
- Rindova, V., Williamson, I., Petkova, A., & Sever, J. (2005). Being good or being known: An empirical examination of the dimensions, antecedents, and consequences of organizational reputation. *Academy of Management Journal*, 48(6), 1033–1049.

- 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.
- Strutt, R. J. (1968). *The Life of John William Strutt: Third Baron Rayleigh*. University of Milwaukee Press.
- Stuart, T., Hoang, H., & Hybels, R. (1999). Interorganizational endorsements and the performance of entrepreneurial ventures. *Administrative Science Quarterly*, 44(2), 315–349.
- Waguespack, D. & Fleming, L. (2009). Scanning the Commons? Evidence on the benefits to startups of participation in open standards development. *Management Science*, 55(2), 210–223.
- 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
Replies	Count of email messages (replies) on all IETF listservs that mention ID within 14 days of initial submission
Email Lists	Count of listservs where ID is mentioned at least once
Revisions	Number of times ID is submitted to IETF
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
Published RFCs	Log one plus sum of RFCs previously published by all ID authors
Authors	Total number of authors listed on draft
Log pages	Log of number of pages in Internet Draft
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
Future RFCs	Log one plus sum of RFCs published in future by all draft authors
Standnards-track	Indicator: Published as a standards-track RFC
Draft Year	Year when draft is first submitted to IETF
Publication Year	Year when draft is Published as RFC

Table 2: Summary Statistics

Variable	Individual IDs			Working Group IDs		
	Mean	S.D.	Obs.	Mean	S.D.	Obs.
Published as RFC	0.072	0.258	4013	0.441	0.497	1405
Replies	2.563	7.111	4013	3.072	7.671	1405
Email Lists	1.615	2.283	4013	5.343	5.425	1405
Revisions	2.110	2.074	4013	4.764	3.734	1405
Et Al Dummy	0.207	0.405	4013	0.228	0.420	1405
WG Chair Author	0.251	0.433	4013	0.447	0.497	1405
Unlisted WG Chair	0.035	0.184	4013	0.060	0.238	1405
Published RFCs	0.751	1.078	4013	1.348	1.212	1405
Authors	2.050	1.186	4013	2.248	1.264	1405
Log pages	2.324	0.967	4013	2.536	1.033	1405
Intl Author	0.402	0.490	4013	0.330	0.470	1405
Multi-sponsor	0.326	0.469	4013	0.514	0.500	1405
Days-to-meeting	3.699	0.641	4013	3.716	0.656	1405
Standards-track <sup>†</sup>	0.328	0.470	534	0.614	0.487	1148
Future RFCs	1.793	0.798	534	2.084	0.854	1148

<sup>†</sup>Sample: RFCs published between 1995 and 2003 (as in Table 5).

Table 3: Identity as a Signal<sup>†</sup>

<b>Linear Probability Models of ID Publication</b>				
Unit of Observation = Internet Draft				
Dependent Variable = Published as RFC				
Sample	All Individual IDs		All WG IDs	
WG Chair Author	0.084** (0.01)	0.040* (0.02)	0.103** (0.03)	0.045 (0.04)
Unlisted WG Chair	-0.066** (0.02)	-0.056* (0.02)	-0.031 (0.07)	-0.011 (0.07)
Et Al Dummy	-0.026** (0.01)	0.006 (0.01)	0.048 (0.04)	0.071 (0.05)
Published RFCs		0.026** (0.01)		0.025+ (0.01)
Log Pages		0.004 (0.00)		0.027* (0.01)
Intl Author		-0.014+ (0.01)		-0.035 (0.03)
Multi-Sponsor		0.004 (0.01)		0.090* (0.05)
Days-to-Meeting		0.038** (0.01)		0.048* (0.02)
Constant	0.058** (0.00)	-0.076* (0.03)	0.385** (0.02)	0.240* (0.10)
Author Count Effects	N	Y	N	Y
IETF Meeting Effects	N	Y	N	Y
Observations	4013	4013	1405	1405
R-squared	0.021	0.045	0.012	0.035
Mean of DV	0.072	0.072	0.441	0.441

Robust standard errors in parentheses; +10% significance; \*5% significance; \*\*1% significance. <sup>†</sup>See Table A3 for additional models that control for rank of Unlisted Chair in list of authors.

Table 4: Identity and Attention<sup>†</sup>

<b>OLS Regressions</b>						
Unit of Observation = Internet Draft						
Sample = Individual Drafts						
Dependent Variable	Replies		Email Lists		Revisions	
WG Chair Author	1.454** (0.33)	1.465** (0.45)	0.784** (0.12)	0.609** (0.12)	0.501** (0.09)	0.292** (0.11)
Unlisted WG Chair	-0.944 (0.63)	-0.750 (0.63)	-0.435+ (0.24)	-0.397+ (0.24)	-0.474** (0.18)	-0.457* (0.18)
Et Al Dummy	-0.006 (0.28)	-0.427 (0.38)	0.111 (0.09)	0.029 (0.12)	-0.098 (0.08)	-0.004 (0.10)
Published RFCs		-0.099 (0.16)		0.113+ (0.06)		0.133** (0.04)
Log Pages		0.023 (0.10)		0.149** (0.04)		0.143** (0.04)
Intl Author		-0.060 (0.24)		-0.021 (0.07)		-0.028 (0.07)
Multi-sponsor		0.363 (0.32)		0.065 (0.12)		0.134 (0.09)
Days-to-Meeting		-0.179 (0.18)		-0.080 (0.06)		0.180** (0.05)
Constant	2.232** (0.13)	1.977** (0.71)	1.411** (0.04)	0.910** (0.24)	2.021** (0.04)	0.836** (0.22)
Author Count Effects	N	Y	N	Y	N	Y
IETF Meeting Effects	N	Y	N	Y	N	Y
Observations	4013	4013	4013	4013	4013	4013
R-squared	0.007	0.022	0.020	0.037	0.011	0.028
Mean of D.V.	2.563	2.563	1.615	1.615	2.110	2.110
$H_0$ : Chair + Unlisted Chair = 0 (p-value)	0.353	0.226	0.103	0.354	0.863	0.325

Robust standard errors in parentheses; +10% significance; \*5% significance; \*\*1% significance.<sup>†</sup>See TableA4 for similar results based on a Poisson specification.

Figure 1: The IETF Publication Process

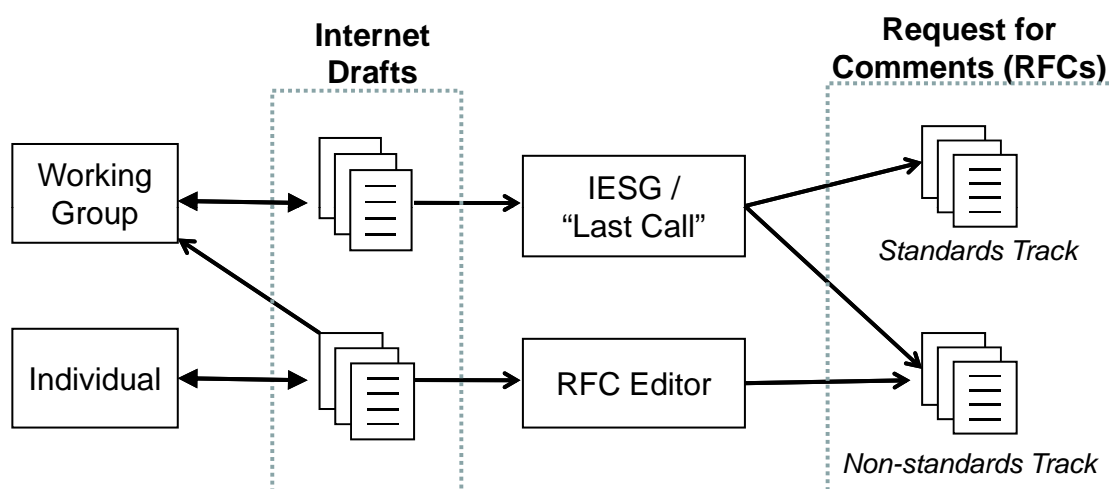


Figure 2: Author Names in ietf-announce and Internet Drafts

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.

<p>A New Internet-Draft is available from the on-line Internet-Drafts directories.</p> <p>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</p> <p>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 ...</p>	
<p>INTERNET DRAFT  18 Feb 2003</p> <p>Basic Network Mobility Support  draft-wakikawa-nemo-basic-00.txt</p> <p>... (BODY TEXT) ...</p> <p>Authors' Addresses</p>	<p>Ryuji Wakikawa  Keisuke Uehara  Koshiro Mitsuya  Thierry Ernst  Keio University and WIDE</p>
<p>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</p> <p>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</p>	<p>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</p> <p>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</p>

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.

<p>A New Internet-Draft is available from the on-line Internet-Drafts directories.</p> <p>Title : Middlebox Communications (MIDCOM) Protocol Managed Objects</p> <p>Author(s) : M. Barnes et al.</p> <p>Filename : draft-barnes-midcom-mib-01.txt</p> <p>Pages : 16</p> <p>Date : 2003-7-1</p> <p>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 ...</p>	
<p>Internet Draft Document: draft-barnes-midcom-mib-01.txt</p> <p>Category: Standards Track Expires: December 2003</p> <p>Middlebox Communications (MIDCOM) Protocol Managed Objects</p> <p>... (BODY TEXT) ...</p> <p>Authors' Address</p> <p>Mary Barnes Nortel Networks 2380 Performance Drive Richardson, TX 75082 USA Phone: 1-972-684-5432 Email: mbarnes@nortelnetworks.com</p> <p>....</p> <p>David Harrington, Co-chair SNMPv3 WG Enterasys Networks</p> <p>....</p>	<p>M. Barnes Nortel Networks Wes Hardaker Sparta D. Harrington Enterasys Networks M. Stiemerling NEC Europe Ltd. June 2003</p>
<p>A New Internet-Draft is available from the on-line Internet-Drafts directories.</p> <p>Title : End-to-End VoIP over MPLS Header Compression</p> <p>Author(s) : J. Ash, B. Goode</p> <p>Filename : draft-ash-e2e-vompls-hdr-compress-01.txt</p> <p>Pages : 0</p> <p>Date : 2003-3-6</p> <p>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...</p>	
	<p>Network Working Group Internet Draft &lt;draft-ash-e2e-vompls-hdr-compress-01.txt&gt; Expiration Date: October 2003</p> <p>Jerry Ash Bur Goode Jim Hand AT&amp;T George Swallow Cisco Systems, Inc. March, 2003</p> <p>End-to-End VoIP over MPLS Header Compression</p> <p>... (BODY TEXT) ...</p> <p>7. Authors' Addresses</p> <p>Jerry Ash AT&amp;T Room MT D5-2A01 200 Laurel Avenue Middletown, NJ 07748, USA Phone: +1 732-420-4578 Email: gash@att.com</p> <p>....</p> <p>George Swallow Cisco Systems, Inc. 250 Apollo Drive Chelmsford, MA 01824 Phone: +1 978 497 8143 Email: swallow@cisco.com</p> <p>...</p>

Figure 4: New Internet Draft Submissions

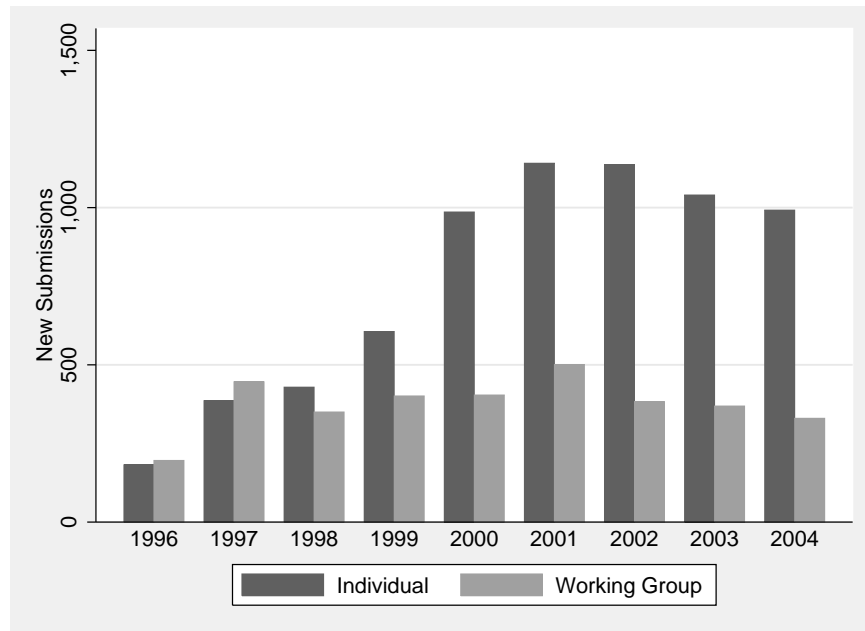


Figure 5: Probability of Et Al

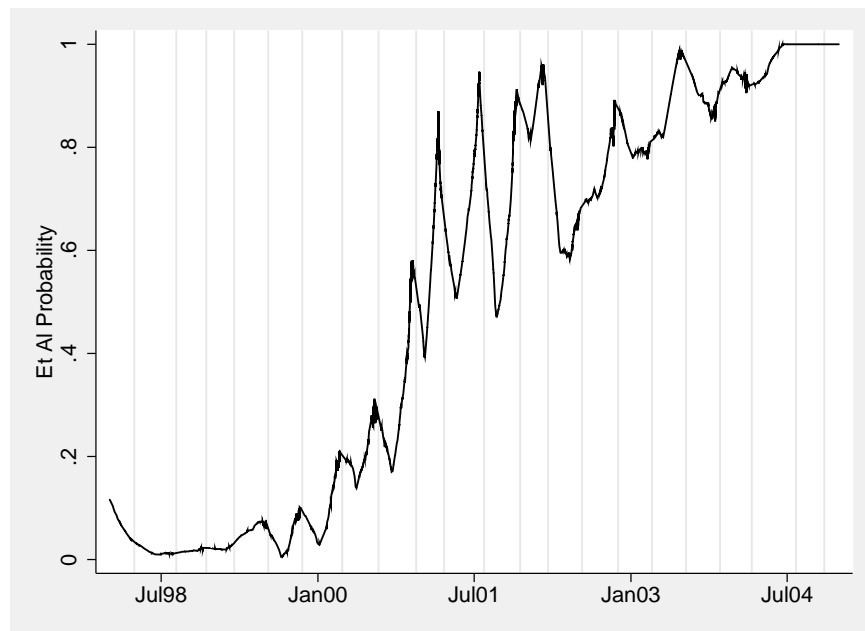
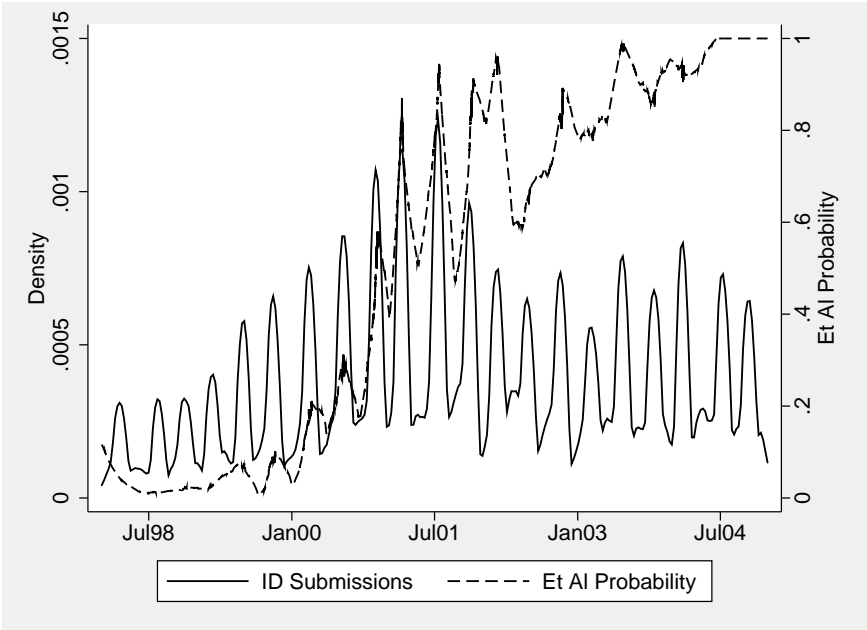


Figure 5 is based on a lowess regression where sample includes all IDs with 2 or more authors. Vertical lines indicate IETF Meeting dates.



Figure 6: Et AI and ID Submission Rates



## Appendix: Robustness Checks

Table A1: Exogeneity of the Et Al Process

<b>Marginal Effects from Logit Rgression</b>				
Unit of Observation = Internet Draft				
Dependent Variable = Et Al Dummy				
Sample	<b>Individual IDs</b> 2+ authors		<b>Working Group IDs</b> 2+ authors	
WG Chair Author	-0.016 (0.03)	0.045 (0.04)	0.007 (0.04)	0.052 (0.04)
Published RFCs		-0.044* (0.02)		-0.058** (0.02)
Multi-Sponsor	0.026 (0.03)	0.038 (0.03)	0.019 (0.05)	0.031 (0.05)
Intl Author	0.014 (0.02)	0.006 (0.02)	0.064+ (0.04)	0.060 (0.04)
Log Pages	-0.003 (0.01)	-0.003 (0.01)	0.010 (0.02)	0.011 (0.02)
Longest Name <sup>†</sup>	-0.005 (0.01)	-0.004 (0.01)	0.009 (0.01)	0.012 (0.01)
Days-to-Meeting	-0.107** (0.02)	-0.103** (0.02)	-0.122** (0.03)	-0.113** (0.03)
Three Authors	0.596** (0.03)	0.596** (0.03)	0.581** (0.05)	0.579** (0.05)
Four Authors	0.794** (0.02)	0.793** (0.02)	0.772** (0.04)	0.775** (0.04)
Five Authors	0.785** (0.02)	0.785** (0.02)	0.802** (0.03)	0.799** (0.03)
Submission Date	0.001** (0.00)	0.001** (0.00)	0.001** (0.00)	0.001** (0.00)
Observations	2256	2256	876	876
$H_0$ : Top-variables jointly zero (p-value)	0.72	0.14	0.24	0.03

Robust standard errors in parentheses; +10% significance; \*5% significance; \*\*1% significance. <sup>†</sup>Number of digits in the last name of author with the longest surname.

Table A2: Status and Reputation

<b>Logistic Regressions</b> <b>(Marginal Effects)</b> Unit of Observation = Internet Draft Dependent Variable = WG Chair Author		
Sample	All Drafts	Single Author
Published RFCs	0.240** (0.01)	0.246** (0.02)
Two Author	-0.075** (0.02)	
Three Author	-0.044* (0.02)	
Four Author	0.011 (0.03)	
Five Author	-0.009 (0.03)	
Year Effects	Y	Y
Observations	4013	1757
Mean of DV	0.19	0.16

Robust standard errors in parentheses; +10% significance; \*5% significance; \*\*1% significance. Sample for column 2 includes only single-authored drafts.

Table A3: Robustness of Signaling Effect

Supplemental Models of Individual ID Publication					
Unit of Observation = Internet Draft					
Dependent Variable = Published as RFC					
Specification	Logit MFX	OLS	OLS	OLS	OLS
Sample	All Drafts	All Drafts	All Drafts	Two or more Authors	2 Authors & No Lead Chair
WG Chair Author	0.081** (0.01)	0.040* (0.02)	0.040* (0.02)	0.044* (0.02)	0.054 (0.03)
Unlisted Chair	-0.032* (0.02)	-0.057* (0.03)		-0.051* (0.03)	-0.059 (0.04)
Et Al Dummy	-0.025** (0.01)	0.007 (0.01)	0.006 (0.01)	-0.002 (0.01)	-0.004 (0.02)
Missing Et Al <sup>†</sup>		-0.002 (0.02)			
Missing Et Al * Chair		0.003 (0.05)			
Unlisted Chair2 <sup>††</sup>			-0.087** (0.02)		
Unlisted Chair3 <sup>††</sup>			-0.034 (0.04)		
Unlisted Chair4 <sup>††</sup>			-0.068+ (0.04)		
Unlisted Chair5 <sup>††</sup>			-0.061+ (0.03)		
Published RFCs		0.026** (0.01)	0.026** (0.01)	0.016* (0.01)	0.006 (0.01)
Log Pages		0.004 (0.00)	0.004 (0.00)	0.005 (0.00)	0.006 (0.01)
Intl Author		-0.014+ (0.01)	-0.014+ (0.01)	-0.015 (0.01)	0.006 (0.01)
Multi-Sponsor		0.004 (0.01)	0.004 (0.01)	0.009 (0.01)	0.008 (0.01)
Days-to-Meeting		0.039** (0.01)	0.039** (0.01)	0.044** (0.01)	0.049** (0.01)
Constant		-0.076* (0.03)	-0.076* (0.03)	-0.117** (0.04)	-0.149** (0.05)
Author Count Effects	N	Y	Y	Y	Y
IETF Meeting Effects	Y	Y	Y	Y	Y
Observations	4013	4013	4013	2256	898
R-squared		0.045	0.045	0.047	0.045

Robust standard errors in parentheses; <sup>+</sup>10% significance; \*5% significance; \*\*1% significance. <sup>†</sup>Missing EtAl = 1 iff the ietf-announce message does not list all authors and does not use *et al*to acknowledge missing names. <sup>††</sup>Unlisted ChairX = 1 iff the  $X^{th}$  author is a WG chair and is not listed in the ietf-announce message.

Table A4: Identity and Attention: Poisson Specification

<b>Poisson Regressions</b>						
Unit of Observation = Internet Draft						
Sample = Individual Drafts						
Dependent Variable	Replies		Email Lists		Revisions	
WG Chair Author	0.502** (0.10)	0.506** (0.15)	0.439** (0.06)	0.344** (0.06)	0.222** (0.04)	0.130** (0.05)
Unlisted WG Chair	-0.296 (0.22)	-0.224 (0.22)	-0.228+ (0.13)	-0.210+ (0.13)	-0.210* (0.09)	-0.204* (0.09)
Etal Dummy	-0.003 (0.11)	-0.169 (0.16)	0.071 (0.06)	0.015 (0.07)	-0.048 (0.04)	-0.003 (0.05)
Published RFCs		-0.038 (0.06)		0.061+ (0.03)		0.059** (0.02)
Log Pages		0.009 (0.04)		0.094** (0.02)		0.069** (0.02)
Intl Author		-0.025 (0.10)		-0.012 (0.04)		-0.014 (0.03)
Multi-sponsor		0.139 (0.12)		0.040 (0.07)		0.067 (0.04)
Days-to-Meeting		-0.068 (0.07)		-0.048 (0.04)		0.085** (0.03)
Constant	0.803** (0.06)	1.431** (0.44)	0.345** (0.03)	0.327 (0.23)	0.703** (0.02)	-0.057 (0.15)
Author Count Effects	N	Y	N	Y	N	Y
IETF Meeting Effects	N	Y	N	Y	N	Y
Observations	4013	4013	4013	4013	4013	4013
Mean of D.V.	2.563	2.563	1.615	1.615	2.110	2.110
$H_0$ : Chair + Unlisted Chair = 0 (p-value)	0.319	0.200	0.072	0.287	0.874	0.381

Robust standard errors in parentheses; +10% significance; \*5% significance; \*\*1% significance.