

# Opening Access to Research

Mark Armstrong  
Oxford

Hamburg: March 2015

# Two claims

- Many publishers charge libraries (let alone the wider public) too much for content
  - perhaps less of an issue for economics than other subjects?
- Many journals force delays on authors and readers with excessive “polishing” of submissions
  - perhaps more of an issue for economics than other subjects?
- Can both problems be addressed with open access regulation?

[paper to appear in the *Economic Journal*]

# Some funding models

- Subscription funded:
  - *Small*: library subscription \$5256
  - bundle of six *AER* journals: subscription \$735
  - average research-oriented university spends about \$6m on journal subscriptions
- Open access:
  - to read “Jealously in dogs” in *PLOS ONE* costs nothing, but author paid \$1350
  - [*PLOS ONE* published 26,000 articles in 2012]
- Delayed open access:
  - *PNAS*: free access after six months
- Hybrid:
  - *Economic Journal*: subscription £477, but author can opt to pay £1500 to make article free to readers

- 1 Certifying article's quality via journal's reputation
- 2 Polishing manuscript with suggestions from editors and referees
- 3 Distributing article to those (and only those) permitted to read it

# Certification

- Authors value certification
  - people working in same area anyway know an author's skill
  - but author cares about other people's evaluation too (tenure committees, job offers, Nobel Prize)
- Readers value certification
  - an article is an "experience good"
  - readers value *ex ante* information about quality before deciding whether to read a paper
- But nowadays other signals in addition to journal's name
  - *Google scholar* citations etc
  - increasingly used in committees (and even by journal editors deciding what to publish)
  - wisdom of crowds vs. a couple of referees (who probably disagree anyway)
  - but danger of inefficient herding?

- Gotten out of hand?
  - Ellison (2002): papers have tripled in length, major revisions now the norm (rather than being “embarrassing”), average length of time in top journals from submission to acceptance is two years
  - McAfee (2010): “we have transformed the business of refereeing from the evaluation of contributions with a little grammatical help into an elaborate system of glacier-paced anonymous co-authorship”
- Also costly to administer
  - “pure certification” service is cheaper for journals, and less onerous for referees and authors
  - and delivers research to readers more quickly

- Authors value readership
  - but care mostly about other researchers (who generate citations) not the “wider public”
- But some benefit to the wider public seeing latest research
  - researchers in poorer countries/institutions
  - vast public interest in medical research
- Putting content behind a paywall does make a difference:
  - big empirical literature (McCabe etc.)
  - bigger difference for downloads than for eventual citations
  - consistent with view that the readers who generate citations have easy access to subscription content

- Internet allows authors to bypass journal distribution:
  - author can post her paper on own webpage, institution's repository or subject-specific database (e.g., *Repec*)
  - can be done with or without permission of publisher
  - authors have incentive to be “careless” about copyright: they do not get subscription income themselves, and benefit when more people see their article
- Propensity to “self-archive” varies across subjects
  - lots in economics (perhaps because of excessive polishing), physics, maths, computer science
  - rare for books in humanities (authors lose out on royalties)
  - medical journals do not allow submission of papers which have previously been posted on the web (“Ingelfinger rule”)
  - even when allowed, many authors do not bother to self-archive, so need element of compulsion in open access policy



# Pros and cons of open access: distribution

- Like other difficult markets, journal market is a “competitive bottleneck” (Armstrong, 2006)
  - authors can publish paper only in a single journal (“single-homing”)
  - each article is a mini-monopoly
  - libraries have to subscribe to wide range of journals to get most of the articles they need (“multi-homing”)
  - outcome often that authors can publish for free, while libraries pay high prices
- See also: mobile call termination, credit cards, newspapers, search engines

# Example

- Initially two parties: author and reader
  - author values reader seeing article at \$10, reader values seeing article at \$10
  - it costs the author some small amount to deliver article to reader (say, posting on internet)
  - joint surplus of interaction, which involves no monetary exchange, is about \$20
- Now third party, a commercial journal, enters
  - makes offer to author to deliver article to reader for free, which author has small incentive to accept
  - journal charges reader \$10 for access
  - joint surplus of author and reader is now \$10, remainder siphoned off by journal

# “Gold” open access policy

- Regulated author must publish in journal which makes article available immediately for free to readers
  - entirely overcomes monopoly pricing problem
  - library spend reduced, wider public has access
- But authors will likely have to pay to publish
  - even with research funds, there's an opportunity cost
  - supply of articles will contract
  - [extreme case: supply of new economics textbooks would fall if authors had to pay rather than be paid]
  - article not being published is even worse for readers than high prices

# “Green” open access policy

- Regulated author must self-archive accepted version or publish in journal with (not too) delayed open access
  - “inferior” version made freely available
  - gives libraries a reasonable outside option, so forces journal to charge less for published version
  - may still leave enough revenue from subscriptions to allow authors to publish for free
  - supply of articles not reduced, so can be better for readers than Gold policy

# Pros and cons of open access: polishing

- With gold policy, subscription price cannot adjust for value-added supplied by journal
  - pretty formatting, valuable comments from referees
  - usually, when a seller provides value-added it can charge a higher price
  - but when price is fixed (at zero), incentive to provide value-added is lower
  - gold policy might stimulate a move to something like a “pure certification” role for journals
- Example:
  - if journal makes authors expend polishing effort  $e$ , its cost per article is  $C(e)$  and its revenue per article is  $R(e)$
  - if author fee is  $p$ , number of authors is  $N(p + e)$
  - journal chooses  $e > 0$  to maximize  $R(e) - C(e) - e$
  - if regulation forced  $R \equiv 0$  then  $e = 0$

# Pros and cons of open access: certification

- The most important part of journal value-added is the certification service
  - might open access regulations mean that less certification in equilibrium?
  - if so, this harms readers, who want guidance in what to read, and (good) authors who want to attract reader and be seen to publish a good article

# Example

- An author has paper which might be good or bad (exogenous)
  - readers in aggregate willing to pay \$10k for article known to be good, nothing for bad article
  - author knows quality, and enjoys some intrinsic benefit from publishing in journal with higher average article quality
  - readers cannot observe quality without effort of reading or journal certification
- Journals come in two forms (many of each)
  - a non-discriminating journal publishes anything, incurs no costs
  - a discriminating journal inspects each submission, which costs \$5k each time (but no other costs), and publishes only good papers
  - whether a journal is discriminating is common knowledge (reputation etc)

# Example

- Without open access regulation:
  - discriminating journals compete for good papers
  - good author publishes for free, readers pay \$10k for access to discriminating journal
  - journal name is perfect signal of quality
  - bad papers go unread
  - it costs an author no more to publish in discriminating journal
- With gold policy:
  - journals must recover cost of peer-review from authors
  - to publish good paper in discriminating journal costs author \$5k
  - only do so if intrinsic benefit from publishing well is greater than \$5k or author has generous research funds



- Some good papers will appear in non-discriminating journals
  - author either has insufficient funds or does not care enough
  - if fraction of good papers in non-discriminating journals is low, readers will not bother sifting through the dross and good papers go unread
  - if fraction higher, readers will read non-discriminating journal, but incur disutility from having to read bad papers
- Model assumes journal name only signal of quality
  - what about *Google scholar* citations etc?
  - but *early* readers can rely only on journal name
  - but if good papers go unread by early readers, they can never “take off”
  - so journal certification may still be important
- Green open access policy allows good authors to publish for free, and so certification continues to be effective

# Conclusions

- Green policy seems superior to no regulation
  - in many subjects (though not economics), little propensity to self-archive
  - indirect impact on the prices publishers can charge to libraries
  - compatible with authors publishing for free
  - need to ensure delays not too long (cf. UK policy)
- [Another policy, not yet on the agenda, is to require regulated authors to publish in “cheap” but not quite free journals]
- Gold policy more radical and disruptive
- But in longer term, move to pure certification could:
  - reduce journal costs, making author fees in gold regime affordable to most
  - reduce delays in getting research to readers
  - make life as an author more enjoyable