# The Job Market's First Steps: Using Research Tools to Simplify the Process

Ryan T. Moore, Washington University in St. Louis and University of California–Berkeley Andrew Reeves, Boston University

Labor Day Annual Meeting of the APSA, when schools start to post their openings for the next year. As we entered the job market, we found ourselves repeatedly collecting information about available positions as we prepared application materials. We monitored APSA's eJobs website, cut and pasted relevant job information into a single spreadsheet, and assembled letters using that information. Here, we introduce free and open-source tools to automate these data collection and letter generation procedures using R and LaTeX. Our system minimizes manual data entry by extracting and creating a spreadsheet from APSA's eJobs information. We walk applicants through the initial job search steps, including using eJobs, compiling position information, and producing attractive letters.

s we entered the job market, we found ourselves spending hours collecting information about job openings and preparing applications to send to hiring committees. First, we repeatedly transferred information about dozens of jobs from the web to a single spreadsheet. This process involved line-by-line cutting and pasting for every position to which we applied (most of which were drawn from APSA's eJobs). Second, we manually created customized letters that drew information from our spreadsheet. As users of LaTeX, a free and open-source platform for creating professionally typeset documents, we found no off-the-shelf mailmerge procedure in LaTeX that accepted a spreadsheet as an input.1 LaTeX is increasingly used by social scientists and taught to graduate students in political science programs because of its flexibility, quality, and affordability (it's free!). We found ourselves wishing we could automate these processes to populate our jobs spreadsheet more quickly and then generate attractive mail-merged letters for potential employers.

To save job-seekers time and effort, we here introduce muRL,<sup>2</sup> a set of tools for collecting job information and preparing cover letters and mailing labels. These tools can be applied to any mailmerge task (e.g., letters of recommendation), and we provide guidance on special methods to simplify job searches in the field of political science. There are several benefits to our approach. We are able to automate the data entry of job information from APSA's

Ryan T. Moore is an assistant professor of political science at Washington University in St. Louis. He is currently a Robert Wood Johnson Foundation Scholar in Health Policy Research at the University of California, Berkeley, and the University of California, San Francisco. He can be reached at rtm@wustl.edu.

Andrew Reeves is an assistant professor of political science at Boston University. His research focuses on elections and political behavior. He can be reached at areeves@bu.edu.

eJobs listings. Using R and LaTeX, we can create handsome documents and minimize the effort dedicated to word processing and formatting tasks, allowing applicants to focus on creating high-quality content to send to hiring committees. Finally, by automating the creation of letters, we can help prevent small mistakes with potentially large consequences: no search committee member from University X wants to read a cover letter that touts an applicant as "a great match for the position at University Y"!

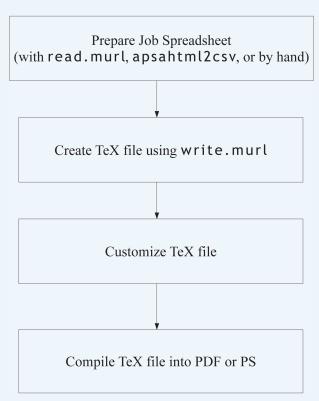
The next section briefly sketches the job market process and the importance of the cover letter. We then detail the benefits of our approach and how it can be used in your own political science job search. Along the way, we introduce the eJobs interface to readers who may be unfamiliar with this tool.

#### **OVERVIEW OF THE JOB MARKET PROCESS**

The process of applying for an academic job follows predictable rhythms, and we advise applicants to review the work of other scholars that thoroughly details these patterns in political science (Carter and Scott 1998; Drezner 1998; Simien 2002). We provide a synopsis here, focusing on the creation of the application packet.

Departments begin posting jobs through APSA's eJobs listing service toward the end of the summer. Applications are generally due between early September and December, with most deadlines occurring in October and November. The eJobs listing is a primary source for U.S. academic jobs, although other sources may be consulted.<sup>3</sup> After selecting a list of potential jobs, the applicant prepares a packet of materials that search committees will consult in making their decisions about which candidates to bring in for interviews and ultimately hire.<sup>4</sup> The typical job market packet consists of a cover letter, a curriculum vitae, three letters of recommendation, and one or two writing samples. As Carter and

Figure 1
The muRL Protocol



Scott have noted, "Without being overly ostentatious, the package should be attractive and effectively organized, and the most relevant and important information that you want to convey should jump out at the reader who just 'skims' the materials"

(1998, 617). The applicant will assemble the packets and mail them to the institutions to which he or she is applying. The administrative task of mailing 10, 20, 50, or even hundreds of these documents can be daunting; muRL tames the process by automating several of the most time-intensive tasks.

#### The Cover Letter

The cover letter is like an abstract of a journal article, but instead of summarizing the arguments and findings of a single piece of research, it highlights and summarizes your career and promise as a political scientist. The cover letter often provides the search committee with its first impression of an applicant, and successful applicants invest significant effort in collecting and presenting its content (Simien 2002). In this letter, you will introduce yourself and provide relevant information such as when you expect to complete your dissertation, what your dissertation is about, the contribution that it makes to the discipline, any publications you have had, and your qualifications as a teacher.

The applicant should carefully consider each department's advertised needs; evaluate how well he or she fits the department, institution, and position; and then convey this match to the search committee in the cover letter. This letter directs the thinking of committee members as they dig deeper into the file and read the letters of recommendation or the writing sample. A committee member may be reviewing well over 100 files, and so an attractive, high-quality cover letter can distinguish an applicant and spark an interest to further investigate his or her file; as King says, "Quality may be everything, but ... style provides important signals" (2006, 122). muRL ensures an attractive presentation and allows the focus to be placed on the quality of your content.

#### SIMPLIFYING THE JOB MARKET WITH RESEARCH TOOLS

Political science training increasingly involves learning software for data processing, statistical estimation, and scientific presentation. Having invested time learning these tools for research

Figure 2
Typical User-Created Spreadsheet of Job Entries

Note. Columns represent fields of job information; rows represent available positions. Columns include the institution name, job type, deadline, contact individual, department and position titles, and mailing information.

purposes, we encourage applicants to exploit them in order to simplify the creation of job packet materials. We employ R to both accelerate the gathering of job data and manage the LaTeX typesetting system to produce a file of ready-to-mail letters.

Why prepare these documents with LaTeX? This free, open-source typesetting system creates beautiful documents so that the user can focus on content and structure. However, until now, LaTeX has been difficult to use for mail-merges, especially when data are being drawn from a spreadsheet. We have simplified and automated this process to allow even novice users to quickly gain LaTeX's typesetting advantages without investing too much time in learning a new way to process documents.

Why do we scrape and process text with R? R is a free language that is well-known for statistical estimation, but it can also collect, organize, and output text as data. For readers unfamiliar with R, a wide array of introductory materials is available (Crawley 2007; Venables, Smith, and the R Devel-

opment Core Team 2010). R enables users to scrape the web and parse text without being required to learn a new general-purpose computing language. In fact, R users can often harness the strengths of other languages such as Python, C++, or Perl through add-on packages. Jackman (2006) offers a succinct demonstration of these capabilities as applied to vote data on the web.<sup>5</sup> Scholars have created general tools that can put R to work scraping web pages and manipulating text (Feinerer 2008; Feinerer, Hornik, and Meyer 2008; Temple Lang 2009a; Temple Lang 2009b). Specific implementations include the automatic collection and processing of U.K. House of Commons voting data (Firth and Spirling 2003; Firth and Spirling 2006).

We provide a suite of functions in R that directly imports APSA eJobs data into a spreadsheet, reads any spreadsheet of job information, and outputs a single file of customized letters and mailing labels. The next sections offer examples of how to use these tools to automate data collection and letter creation, so that less time can be spent preparing documents. Users of R and LaTeX should find the implementation of our approach straightforward; for readers who are new to these tools, the process of obtaining muRL, R, and LaTeX is relatively simple and free.<sup>6</sup> There are many reasons why these free and open-source products are well-suited for conducting analyses and writing papers.<sup>7</sup> Here, we provide an efficient method of conducting mail-merges

Figure 3 **Typical APSA eJobs Search** 

for any purpose, although we focus on political science job application letters.

Figure 1 outlines our method. First, you prepare a spread-sheet of job data, such as that shown in figure 2. This preparation can be done either by hand or automatically by using our apsahtml2csv function (which we detail in the following section). Second, muRL creates a TeX file using write.murl, a function that processes the spreadsheet and incorporates your letter contents. Third, you customize the letter, which may involve altering the TeX file to adjust margins, font sizes, and so on. Last, you compile the TeX file to create a single PDF or PS file containing all the letters.

#### AUTOMATED SCRAPING OF APSA EJOBS DATA

The first step outlined in figure 1 is preparation of the job spreadsheet. Without muRL, this step involves a substantial amount of data entry. For each job, the user must manually enter the address and other job-specific information into a spreadsheet that will later be used for the mail-merge. If you are applying to positions listed in eJobs, our tool automates the data collection and entry processes for you. Even if you choose to forgo the rest of muRL's functionality, its ability to instantly process many eJobs listings into a spreadsheet can save you considerable time.

Figure 4
Typical APSA eJobs Listing

Automatically populating a spreadsheet of APSA eJobs information requires two steps: saving the position information presented in the eJobs listing and processing this information using muRL's apsahtml2csv function. To save the job informa-

tion, first log on to the APSA eJobs webpage. Access to this page requires current APSA membership, a worthwhile investment if you are looking for political science positions. After logging on, locate the jobs that interest you by entering search criteria such as subfield specialty, geographical region, and start date. Figure 3 shows this interface.<sup>8</sup>

When you find a position of interest, navigate to the eJobs summary page for that position. Here, you will find a full description of the position, contact information for the hiring committee, and other related details. These job data will be formatted in the same manner as the example in figure 4.

From your web browser's menu bar, select "Save Page As" or "Save As" and save only the page's source html code into a directory (we recommend saving with the extension ".html"). Then repeat this process for every job that interests you. You will now have a directory of .html pages and need only run a single muRL command to convert this raw information into a spread-sheet. If (a) the directory of

.html pages is on the desktop $^9$  and is labeled "eJobs," (b) the desired output file name is "myjobs.csv," and (c) you saved all the files with extension ".html," then the muRL command will be

 $apsahtml2csv(directory = ``\sim'/Desktop/eJobs", \\ file.name = ``myjobs.csv", \\ file.ext = ``.html")$ 

Of course, you could produce the spreadsheet input to muRL in any way you desire. If you are applying to many political science jobs, however, the apsahtml2csv function will save a great deal of time. Other tools can be combined with our package to create even greater efficiencies in the job application process. For example, you could use apsahtml2csv to create the spreadsheet and then store this information online as a collaborative document using an interface such as Google Documents or Dropbox. Several applicants could split the work of making amendments to the spreadsheet, because a particular position's address, contact person, and so on would be the same for every applicant.<sup>10</sup>



#### FROM THE SPREADSHEET TO THE LETTERS

Once the spreadsheet of job application information has been created either manually or by using the apsahtml2csv function, you can process it by calling read.murl, a single function from our muRL package.<sup>11</sup> If the spreadsheet is named myjobs.csv on the desktop, read.murl processes the data into an R object (here called "murljobs") so that write.murl can process it later<sup>12</sup>:

murljobs <- read.murl ("~/Desktop/myjobs.csv")

Calling zip.plot (murljobs) creates a map of the zip codes in your spreadsheet, allowing you to see the geographic distribution of jobs. Figure 5 displays the locations of the positions from the data in figure 2's spreadsheet.

Next, you can use a single command (write. murl) to create a TeX file. Many options are available to customize the preamble, body, and formatting information to be included in the relevant letters. Full documentation is available for the complete array of options available through arguments in each of muRL's functions. Most important, you should customize your letters to emphasize how your qualifications will benefit the particular department or departments to which you are applying. You can specify the body of the letters directly by using the letter.text argument:

write.murl (murljobs, letter.text

= "This is sample body text.")

Alternatively, you can create the body of your letters in a separate file, using a text editor such as Emacs, Notepad, or TextEdit. If you save this file as content.txt, you can use the letter.file argument to automatically insert its content into your TeX document:

write.murl(murljobs, letter.file = "content.txt")

All that remains is for you to compile the TeX file and send the letters. muRL produces a single file of the letters and associated mailing labels (including postal bar codes). Figure 6 presents an example of a letter created by muRL; the accompanying barcoded mailing labels are shown in figure 7. If you compile the TeX file using pdflatex, a single PDF file will be created. If the search committee asks that letters be submitted electronically, the PDF file can easily be separated into several letters before uploading or e-mailing. This task can be accomplished by using LaTeX's pdf-pages package or other PDF utilities.<sup>13</sup>

Conscientious job seekers are likely to peruse eJobs frequently to stay apprised of the new openings on the market that continually emerge. Prior to mailing applications, we recommend that you regularly add .html pages to the chosen directory and then process the full set of pages using apsahtml2csv. As positions emerge after you have begun sending out packets, you can simply

Figure 6

## **Typical muRL Application Letter**

123 Venus Flytrap Way Cincinnati, OH 45201 jfever@wkrp.edu http://www.wkrp.edu/jfever 513-555-5664

February 6, 2010

Dr. Richard Sanders Graduate Admissions Committee 123 Hello Way Frederick MD 21701

Dear Dr. Sanders:

I will complete my Ph.D. at WKRP University in the Department of Musicology and will defend my dissertation in June 1972.

My dissertation examines ...

In addition to my dissertation, I have also worked on  $\dots$ 

In addition to my research, I believe I could be an an asset to your department through the courses I could offer  $\dots$ 

I have included with this letter, two chapters of my dissertation, my curriculum vitae, an outline of my dissertation, and a summary of my teaching experiences. You will also be receiving letters of recommendation from Professors Less Nessman, Jennifer Marlowe, and Herb Tarlek, which will arrive under separate cover. Thank you for your consideration of my application. I look forward to hearing from you.

Sincerely,

Johnny Fever

create a new directory with the .html pages of the most recent positions and use the function whenever you are ready to apply. Since apsahtml2csv scrapes and stores the application closing date, you may also subset jobs by due date, subspecialty, or any other parameter.

#### **CONCLUSIONS**

Free and open-source tools exist to automatically produce and format many common types of materials, including statistics exams (Grün and Zeileis 2009), quantitative teaching tools (Dryver 2009), analysis output tables (Malecki 2009), and the like. Our package automates the creation of customized letters for job applications, recommendations, or any other purpose. We are also able to automate the collection of data from APSA's eJobs webpages, which constitutes the most comprehensive collection of political science jobs in the United States. Routinizing these tasks saves time and allows you to focus on producing high-quality applications and research content.

#### Figure 7

## Typical muRL Mailing Labels

#### Indianillanillannillalad

DR RICHARD SANDERS GRADUATE ADMISSIONS COMMITTEE 123 HELLO WAY FREDERICK MD 21701

### HaalladdaalHaaalddlad

DOCTOR GARY SANDY GRADUATE ADMISSIONS COMMITTEE 63 NATHAN LANE AUSTIN TX 78702

#### ldallaladdllaaallddd

REV GORDON JUMP GRADUATE ADMISSIONS COMMITTEE 227 MARLA GIBBS RD DETROIT MI 48201

#### Idlandlandladladdadl

PROF FRANK BONNER GRADUATE ADMISSIONS COMMITTEE 348 LB LANE DEPT OF ROCK MUSIC PEORIA IL 63112

#### Idhadhadhallaadddal

MISTER HOWARD HESSEMAN GRADUATE ADMISSIONS COMMITTEE 21 RECURSIVE LOOP DEPT OF NYC TOPEKA KS 66603

#### Haladlalladlaadlal

DR LONI ANDERSON GRADUATE ADMISSIONS COMMITTEE DEPT OF POLITICAL SCIENCE JACKSON WY 83001

We have also used these tools to produce letters of recommendation. We ask our students to provide us with a muRL-ready spreadsheet of position data to streamline the recommendation letter-writing process. This approach allows us spend more time preparing compelling evaluations of our students and less time cutting and pasting.

Creating application materials is a critical initial step in political science job searches. Job seekers may present themselves awkwardly, waste time on repetitive tasks, and make small but embarrassing mistakes during this process. Here, we offer flexible tools tailored to help applicants avoid these pitfalls.

#### NOTES

We thank Keith Schnakenberg for helpful comments and suggestions on this article.

- 1. The Comprehensive TeX Archive Network lists three other documented approaches: mailmerge (Frasson 2009), mailing (Braams 1994), and datatool (Talbot 2009). However, Frasson's (2009) and Braams' (1994) approaches do not accept rectangular spreadsheet data, which is probably the most natural and user-friendly way to store, organize, and view job information. Our solution accepts this kind of data, requires no new LaTeX style file or package installation, and creates a TeX file directly for the user instead of requiring him or her to declare fields, repetitions, and entries within a user-created TeX document. These merits, implemented through R, minimize the amount of noncontent post-processing required of the user. Talbot's (2009) approach uses a LaTeX package that rests on a database solution. In his own words, "attempting to use ... [LaTeX as a database management system] is like trying to fasten a screw with a knife instead of a screwdriver: it can be done, but it requires great care" (30).
- 2. Software for creating  $m \text{ail-merges} \ u \text{sing} \ R$  and LaTeX.
- 3. Political science job seekers should certainly consult APSA's eJobs listings, as well as other resources such as the Chronicle of Higher Education's job listings, online databases such as http://www.higheredjobs.com/ and http://www.jobs.ac.uk, and their advisors. One important warning: although a number of post-doctoral fellowships are listed on APSA's eJobs, many are not. We used eJobs for all but a handful of our own applications.
- 4. Applicants should tailor their materials to the positions for which they are applying; sound advice on this subject has been written regarding applications for jobs at liberal arts colleges (Donaldson Deardorff et al. 2001) and community colleges (Douglas 2002), as well as for American university employment in general (Nerad 2004; Nelson 1995). Pfannestiel (1998) has described the trend toward temporary, non-tenure track positions.
- R is extremely flexible and has rich input/output capabilities, as well as the ability to process textual data through functions like substring(), gsub(), readLines(), and so on.

- 6. R can be downloaded at http://www.r-project.org/ and is available for Windows, Mac OS X, and Linux platforms. Further information about obtaining R and LaTeX is available at many places on the web, including our own web pages. Once you have R and are connected to the Internet, two R commands install and prepare muRL: install.packages ("muRL"), followed by library (muRL).
- We leave more detailed arguments for the value of these open-source products to others in the academy and the popular press. On LaTeX, see, for example, Goossens, Mittelbach, and Samarin (1994); Lamport (1986); and King (2006). On R, see Vance (2009).
- 8. We provide more in-depth instructions for this section at http://rtm.wustl.edu/files/res/murlejobsPS.pdf.
- We show directory structures here for Mac OS; path name formats will differ slightly for other operating systems.
- 10. The authors maintain the package and update it to reflect occasional changes in eJobs formatting. Additionally, muRL's open-source licensing allows any user to change the code to their liking. See http://www.gnu.org/licenses/gpl.html for more information on the license.
- We provide more in-depth instructions for this section at http://rtm.wustl.edu/files/res/murlvignPS.pdf
- 12. To load the data from figure 2 and follow along with the steps outlined later, simply type data (murljobs) at the R prompt.
- For example, the free tools CutePDF Writer (for PCs) and Skim (for Macs) can save individual letters using short dialogues ("Print > [Select pages] > PDF" in CutePDF Writer; "Save As PDF" in Skim).

#### REFERENCES

- Braams, Johannes L. 1994. Mailing: Create a Mailing. LaTeX package version 1.0. http://tug.ctan.org/tex-archive/macros/latex/contrib/mailing/.
- Carter, Ralph G., and James M. Scott. 1998. "Navigating the Academic Job Market Minefield." *PS: Political Science and Politics* 31 (3): 615–22.
- Crawley, Michael J. 2007. The R Book. Hoboken, NJ: John Wiley and Sons.
- Donaldson Deardorff, Michelle, Marianne Githens, Glen Halva-Neubauer, William Hudson, Grant Reeher, and Ronald Seyb. 2001. "Everything You Always Wanted to Know about Getting and Keeping a Job at a Private Liberal Arts College, but Your Graduate Advisor Didn't Tell You." *PS: Political Science and Politics* 34 (4): 856–57.
- Douglas, Donald Kent. 2002. "Tenure-Track Employment Opportunities at the Community College Level: A View from the Job Candidate's Perspective." *PS: Political Science and Politics* 35 (1): 95–100.
- Drezner, Daniel W. 1998. "So You Want to Get a Tenure-Track Job . . ." PS: Political Science and Politics 31 (3): 609–14.
- Dryver, Arthur. 2009. "The Enhancement of Teaching Materials for Applied Statistics Courses by Combining Random Number Generation and Portable Document Format Files via LaTeX." Journal of Statistical Software 31 (3): 1–9.
- Feinerer, Ingo. 2008. "An Introduction to Text Mining in R." R News 8 (2): 19–22.
- Feinerer, Ingo, Kurt Hornik, and David Meyer. 2008. "Text Mining Infrastructure in R." *Journal of Statistical Software* 25 (5): 1–54.
- Firth, David, and Arthur Spirling. 2003. "Divisions of the United Kingdom House of Commons, from 1992 to 2003 and Beyond." Working paper. http://www2.warwick.ac.uk/fac/sci/statistics/staff/academic-research/firth/software/tapir/firth-spirling.pdf.
- ——. 2006. "tapiR and The Public Whip: Resources for Westminster Voting." Political Methodologist 14 (2): 2–4.
- Frasson, Miguel V. S. 2009. MAILMERGE: Repeat Amount of Text Merging Fields. LaTeX package version 1.0. http://tug.ctan.org/tex-archive/macros/latex/contrib/mailmerge/.
- Goossens, Michel, Frank Mittelbach, and Alexander Samarin. 1994. *The LaTeX Companion*. New York: Addison-Wesley.
- Grün, Bettina, and Achim Zeileis. 2009. "Automatic Generation of Exams in R." Journal of Statistical Software 29 (10): 1–14.
- Jackman, Simon. 2006. "Data from the Web into R." Political Methodologist 14 (2): 11–15.
- King, Gary. 2006. "Publication, Publication." PS: Political Science and Politics 39: 119–25.

- Lamport, Leslie. 1986. LaTeX: A Document Preparation System. 5th ed. Reading, MA: Addison-Wesley.
- Malecki, Michael. 2009. Apsrtable: Apsrtable Model-Output Formatter for Social Science. R package version 0.7-6. http://CRAN.R-project.org/package=apsrtable.
- Nelson, Cary. 1995. "Lessons from the Job Wars: Late Capitalism Arrives on Campus." Social Text 13 (3): 119–34.
- Nerad, Maresi. 2004. "The Ph.D. in the U.S.: Criticisms, Facts, and Remedies." Higher Education Policy 17 (2): 183–99.
- Pfannestiel, Todd. 1998. "It's Not Just a Job, It's an Indenture: Graduate Students and the Academic Job Market." *Academic Bulletin of the AAUP* 84: 44–47.
- Simien, Evelyn M. 2002. "On the Market: Strategies for the Successful Job Candidate." *PS: Political Science and Politics* 35 (3): 581–83.

- Talbot, Nicola L. C. 2009. Datatool: Databases and Data Manipulation. LaTeX package version 2.03. http://tug.ctan.org/tex-archive/macros/latex/contrib/datatool/.
- Temple Lang, Duncan. 2009a. RCurl: General Network (HTTP/FTP/...) Client Interface for R. R package version 1.2-o. http://CRAN.R-project.org/package=RCurl.
- . 2009b. XML: Tools for Parsing and Generating XML within R and S-Plus. R package version 2.6-o. http://CRAN.R-project.org/package=XML
- Vance, Ashlee. 2009. "Data Analysts Captivated by R's Power." *New York Times*, January 7, B6.
- Venables, William N., David M. Smith, and the R Development Core Team. 2010. An Introduction to R. United Kingdom: Network Theory Limited.

CAMBRIDGE

## **JOURNALS**

## International Theory Duncan Snidal, University of Chicago Alexander Wendt, The Ohio State University

**EDITORS:** 

## A Journal of International Politics, Law and Philosophy



**International Theory** promotes theoretical scholarship about the positive, legal, and normative aspects of world politics respectively. IT is open to theory of absolutely all varieties and from all disciplines, provided it addresses problems of politics, broadly defined and pertains to the international. IT welcomes scholarship that uses evidence from the real world to advance theoretical arguments. However, **IT** is intended as a forum where scholars can develop theoretical arguments in depth without an expectation of extensive empirical analysis.

IT's over-arching goal is to promote communication and engagement across theoretical and disciplinary traditions. IT puts a premium on contributors' ability to reach as broad an audience as possible, both in the questions they engage and in their accessibility to other approaches. This might be done by

addressing problems that can only be understood by combining multiple disciplinary discourses, like institutional design, or practical ethics; or by addressing phenomena that have broad ramifications, like civilizing processes in world politics, or the evolution of environmental norms. IT is also open to work that remains within one scholarly tradition, but in that case authors must make clear the horizon of their arguments in relation to other theoretical approaches.

All articles are peer-reviewed.

#### FREE email alerts.

Keep up-to-date with new material. Sign up at: journals.cambridge.org/int-alerts

Recommend International Theory directly from its homepage -FREE online access for you when your library subscribes.

journals.cambridge.org/int

