Supplementary Online Appendix

1. AMS Mathematical Reviews

AMS Mathematical Reviews records the titles, publication sources, author names, references, and citations of over 2 million articles from 2,764 different journals and publication sources worldwide. The AMS team painstakingly assigns the correct set of articles to each person (even in difficult cases such as identical names), tagging each author with a unique author identifier. A staff of professional mathematicians at AMS also records subject classification codes for each paper in its database: 64 subfields defined according to the 2-digit 2010 Mathematics Subject Classification (MSC2010). Beginning in the early 1980s, the archival record of each published paper provides information not only on the author and the field of the paper, but also on the author's affiliation at the time the paper was published. This institutional information, however, is relatively complete only for papers published beginning in 1984.

The AMS provided us with a spreadsheet containing the publication history of all mathematicians since 1900. A row in this spreadsheet defines a particular permutation of author \((i)\), field \((f)\), and year \((t)\). For each \((i, f, t)\) row we have information on: the number of papers published, the total number of citations received by those papers as of 2011; the institutional affiliation(s) associated with the papers in that row; and the (country) location of the affiliation(s). For about 95 percent of the Soviet authors, the AMS also provided a database of information at the publication level, containing: the title, journal, language, and year of the publication, as well as the unique author identifiers of each author of the publication.

Although the count of published papers in these data is complete, the citation information is not. In particular, the AMS database only reports the citations that a paper has received after 2000 (regardless of when the paper was published). In addition, the citation count is based on references made to a paper in a limited list of journals (chosen by the AMS).

We treat co-authorships in two alternate ways. First, regardless of how many authors are responsible for a single paper, each author is given full credit when we count a
mathematician’s number of papers and citations. Second, we give each author \((1/n)\) share of a paper in which there are \(n\) authors.

2. Thomson Reuters’ ISI Web of Science

The ISI Web of Science records the titles, publication source, author names, references, and citations of millions of articles from thousands of journals worldwide. For many articles (especially after 1978), the database records research addresses and reprint addresses for each author, as well as abstracts, keywords, and funding information.

Most important, the database contains complete citation information for each article in a primary set of 7,621 journals, selected to include all of the most important journals in each field. Articles in marginal journals also appear in the database if they either cite an article in the primary database or are cited by such an article. Thus, the ISI Web of Science citation database only misses articles that are not in important journals, are not cited by any articles in important journals, and do not themselves even cite any articles in important journals. This means the citation-adjusted weight of any missed articles must be very small, especially since ISI adds important journals to its main database by subfields, so even marginal subfields likely have the main field journals represented.

We purchased the records of all 1,179,787 articles in the primary ISI Web of Science database between 1970 and 2009 for the following categories: Mathematics, Applied Mathematics, Interdisciplinary Applications of Mathematics, Mathematical Physics, and Statistics & Probability. We also purchased all 1,921,587 articles referenced by these main articles, and all the 2,368,123 papers that cite these main articles.

We obtained special permission from the AMS to merge our ISI papers by title, source, and author with the AMS database. We obtained 882,088 matches out of the 1,753,148 journal articles in the AMS database, or just slightly over a 50 percent match rate. The merging allows us to allocate the number of citations to a paper in the ISI database to the matched paper in the AMS database. In addition, the match also allows us to obtain the location and institutional information for additional papers. In particular, the AMS database reports location information on a systematic basis beginning in 1984, while the ISI database provides the same information roughly since 1978.
3. Mathematics Genealogy Project (MGP)

For every doctoral degree ever awarded in mathematics, the MGP database reports the name of the degree recipient, the name of the doctoral advisor, the year in which the degree was completed, and the name of the institution granting the degree. In addition to these data, the MGP also provided us with each advisor’s and student’s AMS unique author identifier. The identifier allows us to merge MGP database with the AMS/ISI databases, and obtain all the available information on the publication and citation records of both students and advisors. The unique AMS identifiers are available for about 65 percent of degrees awarded since 1960.

4. Construction of the sample of Soviet mathematicians

We focus on a sample of mathematicians whose first published paper occurred on or after 1940, who have fewer than 60 years of “experience” (defined as current year minus the year of first publication), and who were mathematically active prior to the collapse of the Soviet Union (defined as having published at least one paper between 1970 and 1989).

The AMS database reports the mathematician’s institutional affiliation and location for any year that the mathematician published at least one paper. We used the AMS data to identify the number of times that the reported institutional affiliation prior to 1989 was located in the former USSR. The AMS database, however, only began to systematically collect the information on the affiliation associated with any given publication around 1984. We supplement the AMS data by using both the ISI and the MGP data to define the sample of Soviet mathematicians. As noted above, the ISI data contains affiliation and location information for many other published papers between 1978 and 1983. We include this location information in the time series data for any given mathematician if the equivalent data are not available in the AMS.

The affiliation/location information are sometimes missing in both the AMS and ISI databases (even though the author has published a paper in a particular year). In defining the relative number of times that an author uses a Soviet affiliation, we only include years in which the location information is available.

Finally, the MGP database reports the institution where the mathematician received the doctoral degree. A mathematician is then classified as Soviet if the AMS/ISI location
data indicate that at least half of the affiliations of publication prior to 1989 were located in the former USSR. If the pooled AMS/ISI data do not provide any information on the location of the mathematician prior to 1989, we include mathematicians who received a doctoral degree in the Soviet Union between 1950 and 1977 in the universe of Soviet mathematician.

To determine the “émigré” status of a Soviet mathematician, we must observe the mathematician publishing at least one paper between 1992 and 2008. A Soviet mathematician is an émigré if the modal affiliation of publication between 1992 and 2008 is located outside the geographic boundaries of the former Soviet Union and is a “valid” country code in the sense that it provides an actual location for an institution.

5. Construction of the list of top international journals

We first assemble a list of the 20 journals with the highest impact factors in the 2012 Thompson Reuters ISI Journal Citation Reports for "Mathematics", the 20 journals with the highest impact factors in "Physics, Multidisciplinary" (this includes general interest physics journals), the 10 journals with the highest impact factors in "Mathematics, Applied", the 10 journals with the highest impact factors in "Physics, Mathematical", and the 10 journals with the highest impact factors in "Physics, Particles and Fields." We then drop the journals which are expository, publish biographical material, or publish chemical or physical constants. Finally, we add two journals that are too broad to be naturally included in any of the above categories but occasionally publish mathematical results in the above fields: Science and Nature.

The resulting list includes the following 38 journals: Acta Mathematica; Advances in Mathematics; American Journal of Mathematics; Annals of Physics; Annals of Mathematics; Applied and Computational Harmonic Analysis; Chaos; Communications on Pure and Applied Mathematics; Communications in Applied Mathematics and Computational Science; Communications in Nonlinear Science and Numerical Simulation; Computer Physics Communications; Duke Math Journal; European Physical Journal C; Europhysics Letters; General Relativity and Gravitation; Geometric and Functional Analysis; Publications Mathématiques de l'IHÉS; Inventiones Mathematicae; Inverse Problems; Journal of the American Mathematical Society; Journal of Computational Physics; Journal of Cosmology and Astroparticle Physics; Journal of the European Mathematical Society; Journal of High Energy