The State of Replication Code Availability: Evidence from the German Socio-Economic Panel

Lukas Fink*, Jan Marcus†

DRAFT of January 24, 2024! Please do not cite

Abstract

Reproducibility of scientific results is an essential part of the scientific method, and providing replication code, such as scripts in statistical software like Stata, R, or Matlab, is an effective and inexpensive way to facilitate reproducibility. However, little is known about the extent of replication code provision, over time and across disciplines. This study examines the availability of replication code for over 2,500 peer-reviewed articles based on the German Socio-Economic Panel (SOEP), one of the most widely used datasets in economics and other social sciences. The empirical findings reveal that only 6% of SOEP-based studies have available code. However, the share of studies providing code has steadily risen since the early 2010s, reaching almost 20% in recent years. Notably, studies published in high-ranked journals exhibit higher code availability, which aligns with the trend that such journals often enforce mandatory code-sharing policies. Our findings underscore the evolving landscape of code availability and its implications for the transparency and reproducibility of scientific research.

JEL: A12, C80, C88

Keywords: Reproducibility; replication code; code availability; journal policies; SOEP

*Lukas Fink, Freie Universität Berlin, e-mail: l.fink@fu-berlin.de.

†Jan Marcus, Freie Universität Berlin, IZA Bonn, Berlin School of Economics, e-mail: jan.marcus@fu-berlin.de.

1. Introduction

Reproducibility of scientific results is an essential part of the scientific method. It ensures that research can be independently verified and promotes transparency and trust in scientific results. Reproducibility is particularly important in economics because the results of economic studies are used to inform policy decisions and, hence, may have direct consequences for the lives of many individuals. When a study is reproducible, it means that other researchers, using the same data and the same code (i.e. computer instructions), get the same results as the original study (**?**). Hence, the reproducibility depends on two crucial ingredients: the availability of code and the availability of data.

The practice of making code and data available has numerous advantages for the scientific community in addition to allowing for an easy reproduction of the published results. First, it enables others to build upon the work, fostering further investigation and a more comprehensive understanding of the subject matter. Second, errors in the code can be more easily found and corrected, strengthening the self-correction feature of the scientific system.[1] Third, it facilitates knowledge dissemination as others can gain deeper insights into the applied methods from the provided code. Related to this, others can improve their coding skills by learning from the provided code (e.g., with respect to the construction of specific tables and graphs). Fourth, it might serve as a deterrent against fraud and questionable research practices like p-hacking as researchers know that their code might be exposed to scrutiny. Fifth, it can serve as signal that the researchers did not engage in questionable research practices.

Despite the advantages, it is unclear to what extent economists and researchers from other social sciences make data and code available. We aim to contribute to filling this gap by examining the extent of code provision. Hence, we focus on the ingredient for reproducibility that is arguably easier to achieve: Basically every empirical researcher can share code but not everyone can share data as researchers are frequently prohibited from sharing analyzed data due to legal and confidentiality reasons. This is especially relevant in economics, where secondary analysis of data collected by others, such as governmental organizations or general interest surveys, is more common than analysis of self-collected data. Further, providing code is comparatively inexpensive in terms of both time and money for researchers, as it typically takes less time than sharing data.[2] For these reasons, we focus on a setting where data are

[1] For example, Foote and Goetz (2008) found a smaller effect size in replicating Donohue and Levitt (2001)'s study on legalized abortion and crime due to a coding error, while Herndon et al. (2013) identified a contradiction in Reinhart and Rogoff (2010)'s original study on high public debt and low growth rates after correcting an error in an Excel spreadsheet and using different weighting. While these examples of coding errors received a lot of attention, (unintentional) coding errors often have only minor consequences for the conclusions of an article (Laurinavichyute et al. 2022).

[2] For data sharing, researchers may need to label variables and variable labels, and carefully consider which

available to researchers worldwide at no cost, making the provision of code the critical factor for reproducibility.

To get a better understanding of the extent of code sharing and its development over time, this metascientific study focuses on code availability for publications based on one of the most-widely used data sets in economics and other social sciences, the German Socio-Economic Panel (SOEP). For this purpose, we conducted an intensive search for code availability of all SOEP-based publications in peer-reviewed journals in the 1985-2021 period. We checked the journal website for each of the 2,518 articles, the actual articles, the websites of 2,774 unique authors that wrote the articles, and specific online repositories. The collection of this variable is a major contribution of this paper. We use this information to examine different correlates of code provision, the development of code provision over time, and the relationship between code provision and quality metrics like journal impact factors and citations. Further, we study several drivers for the development in code availability over time. For this purpose, we gathered information on journal policies on code sharing from the Transparency and Openness Promotion (TOP) framework (Nosek et al. 2015), journal websites and a survey among editors of almost 100 journals in different disciplines.

The SOEP is particularly useful to study code provision for several reasons. First, the SOEP data are widely used by the scientific community. The SOEP has about 3,500 users worldwide (Goebel et al. 2019) and more than 2,500 peer-reviewed, SOEP-based journal articles have been published. Second, the SOEP data are not only employed by economists but also by researchers from several other disciplines (in particular, sociology, psychology, political science, demography), allowing to examine code sharing behavior in different disciplines. Third, the SOEP group provides a list of all publications that are based on the SOEP data, which allows to clearly define our target population. Fourth, there is a clear separation of the data collector (a fieldwork agency on behalf of the SOEP group at DIW Berlin) from the data analyst. This means that the raw SOEP data can be seen as exogenous to the researchers analyzing the data and that strategic decisions in the data collection process are of little relevance. Fifth, the SOEP data can be downloaded and used free of charge by researchers around the world.[3] This as a major advantage of the SOEP (and similar survey

variables should not be released. This may be to reduce the risk of de-anonymization of specific observations or because they plan to publish another paper using the same data and variables that have not been used.

[3]More specifically, researchers at universities and research institutions can use the data for research and teaching purposes by signing a data distribution contract, wherein they agree to comply with the data security and privacy requirements of the European General Data Protection Regulation. Data protection rules differ inside and outside the EU. Within the EU, users have full access to the SOEP core dataset, outside the EU, they have access to a 95% sample of this dataset. The core SOEP data contain information on the federal state. More detailed regional data are available via remote or on-site access. See SOEP Group (2024) for more details.

data) for studies on replicability and reproducibility, in particular compared to administrative data, which often can only be analyzed in restricted access environments and often only after payment for data access.

Our empirical results show that the share of peer-reviewed SOEP studies with code available is only 6 %. Examining correlates of code availability, we find that single-authored articles and articles written in languages other than English are less likely to provide code. In addition, code availability is positively correlated with higher journal impact factors and more citations. We also see that code provision is less common for studies in economics journals in our sample. When we control for year of publication, the correlation between code availability and article language disappears and the difference between economics and other disciplines becomes smaller, suggesting that parts of these observed correlations are driven by time effects.

We also document that the share of SOEP studies that provide code has increased sharply since the early 2010s, reaching a share of 17.8 % in 2021. We argue that the increase in code provision is driven by a mix of technological advances, individual researcher awareness, and journal policies. Technological advances, including the widespread use of individual researcher websites, the establishment of journal websites, and the creation of specialized online repositories, have reduced the financial and logistical barriers to code sharing. We also present suggestive evidence that top-down initiatives, such as journal policies on code sharing, have played an important role in increasing code availability over time. Moreover, the fact that a substantial fraction of articles only provide information about code availability on the authors' websites, but not on the journal website or in the article, suggests that individual researchers' awareness of the importance of code sharing also plays an important role. It also highlights the need to search the authors' websites in order to get a comprehensive overview of code availability.

Our study contributes to the literature on research transparency and open science practices. While this is a rather broad and interdisciplinary field, we particularly contribute to two strands of literature in this overall field. The first is the state of research transparency practices and factors that are associated with it. Colliard et al. (2023) find that less than 5% of papers published between 2010 and 2020 in one leading journal in finance, the *Journal of Financial Economics*, provide some data or code on the journal's website. Complementing this research, our approach involves distinguishing between the provision of data and code, as well as actively checking the personal websites of authors and data repositories.

Key (2016) studies the availability of data and code in empirical articles published in six leading political science journals in 2013/14. The availability of code and data was examined by following the instructions for replication materials given in the article; authors' personal websites were checked only when the article stated that material was available on an author's website. 58% of the 494 articles contain both data and code, and a further 2% of articles provide code but no data, giving a total of 60% of articles containing replication code, with considerable variation across journals. Key (2016) identifies journal policies as an important predictor of code and data availability. In particular, journals where the provision of replication material is mandatory had significantly higher rates of code and data availability than journals where this provision was only expected or not required. The second strand relates to the development of research transparency practices over time. Here, a general finding in the literature is that different practices of research transparency are on the rise (for an overview on the development in economics, see Miguel 2021). For instance, Miguel (2021) presents evidence that pre-registration and pre-analysis plans[4] are increasingly used in experimental economics and that more and more studies register every year at the Randomized Control Trial Registry of the American Economic Association. Moreover, more and more journals in different disciplines allow for "pre-results review", also known as "registered reports", where a research project is refereed based on its research plan and methodology before the research is conducted and results are obtained (Hardwicke and Ioannidis 2018; Miguel 2021).

Further, journals have implemented various policies to foster research transparencies. For instance, several journals introduced editorial statements underscoring the importance of disseminating findings with non-significant or null results and there is some evidence that these statements increased the share of non-significant results Blanco-Perez and Brodeur (2020). Additionally, many journals implemented explicit code and data sharing policies.

The manuscript is organized as follows: Section 2 describes our data and variables and presents descriptive statistics. Section 3 presents empirical results on the state of replication code availability, its correlates, and its evolution over time, as well as the relationship between code availability and journal impact factors and article citations. Section 4 concludes and provides an outlook for future research.

2. Data

2.1. Basic data: SOEP and SOEPlit

The German Socio-Economic Panel (SOEP) is one of the largest and longest-running household panel surveys in the world: The first interviews were conducted about 40 years ago (in 1984) and currently about 30,000 individuals in about 15,000 households are interviewed annually. The SOEP covers a wide range of topics, including questions on employment,

[4] The purpose of pre-registrations and pre-analysis plans is to enhance confidence in a study's findings by reducing researchers' degree of freedom and the capacity to selectively choose results.

income, health, education, demography, income, housing, life satisfaction, attitudes, values and personality. For all these reasons, it is not surprising that the SOEP is highly regarded by economists and researchers from other social science disciplines: SOEP has more than 3,500 users worldwide, and in the last decade more than 100 peer-reviewed articles have been published each year using SOEP data - and between 250 and 300 other publications (including working papers, books, policy reports, etc.) (Goebel et al. 2019).

Our main dataset is based on SOEPlit, a database that aims to cover all publications based on SOEP data.[5] The SOEPlit database contains bibliographic information about each publication, including title, type of publication, year of publication, language, journal (if applicable), digital object identifier (DOI, if available), and the names of the authors. Our version of SOEPlit is from June 2022 and contains SOEP-based publications up to the year 2021. We focus our analyses on peer-reviewed articles published in journals listed in one of the Clarivate Analytics citation indices, i.e. the Social Science Citation Index (SSCI), the Science Citation Index (SCI), the Arts and Humanities Citation Index (AHCI) and the Emerging Sources Citation Index (ESCI).[6] We do not include working papers for two reasons. First, we want to avoid double counting, since many working papers end up as peer-reviewed articles. Second, authors may be more reluctant to share working paper code for fear of intellectual property theft. In total, our database consists of 2,518 unique peer-reviewed publications.

2.2. Data on code availability

We then checked each publication for publicly available replication code. This was done in four different ways: (i) we examined the journal webpage of each article for code availability; (ii) we searched the authors' webpages for code availability;[7] (iii) We searched specific online repositories for code availability;[8] (iv) We reviewed all 2,518 articles for references to publicly available code.[9]

[5] SOEPlit is maintained by the SOEP group at the German Institute for Economic Research (DIW Berlin).

[6] In our sample, the SSCI accounts for the majority of articles (86%), while the SCI makes up another 10%.

[7] More specifically, our student assistants checked the personal webpages or, if not found, the institutional webpages of the first four authors for code. This covers 2,774 unique authors. Most papers have a maximum of three authors; however, there are few studies with very many authors.

[8] These include the GESIS data archive, Harvard Dataverse, Open Science Framework, openICPSR and Zenodo. We used the search terms "SOEP", "GSOEP", "Socio-Economic Panel", and "Sozio-oekonomisches Panel".

[9] In particular, student assistants reviewed the acknowledgments and searched the paper for key terms such as code, replication, syntax, and Stata.

2.3. Additional data

We integrated additional information at the article and journal level into our dataset. Article-level information includes citations in Google Scholar, obtained through web scraping using the Scholarly package in Python (Cholewiak et al. 2021).[10]

Journal-level information refers to Journal Impact Factor (JIF) metrics as well as the primary discipline of the journal. We collected the two-year and five-year JIF metrics since 1997 from Clarivate's Master Journal List.The JIF is merged by journal name and year of publication.[11] To assign the primary discipline of each journal, we use the classification of Science-Metrix, which categorizes journals into 174 mutually exclusive subfields, nested within 20 mutually exclusive fields (Archambault et al. 2011). We have grouped the subfields into 5 categories: Economics, Sociology, Psychology, Other Social Sciences, and Health & Other Sciences.[12] Appendix Table A1 provides, for each of our five disciplinary categories, the names of the journals with the most articles (as well as the number of articles) in our data set. We merged the journal information from Science-Metrix based on the International Standard Serial Number (ISSN) of the journal.[13]

2.4. Information on journal policies

Additionally, we included information about journals' policies on replication code, following the classification of the Transparency and Openness Promotion (TOP) framework (Nosek

[10]Citations are as of XXXdata

[11]Since the JIF information is only available from 1997 onwards, we impute the impact factor for earlier years based on the 1997 impact factor (this applies to 14% of the articles in our samples). In addition, two-year impact factors are not available for some journals that have only recently been included in the citation indices. This applies to less than 4% of the articles in our sample. We exclude these articles in the regression analyses with JIF as the dependent variable (see Figure 6 and Table A2), while we include a missing dummy in the regressions where JIF is a control variable (Table A3).

[12]Specifically, we have assigned the following subfields of "Economics & Business" to the discipline of "Economics": "Economics", "Econometrics", "Economic Theory", "Finance", "Agricultural Economics & Policy", "Development Studies", and "Industrial Relations", as well as the subfield "Health Policy & Services" of the field "Public Health & Health Services" (since in our sample this consists mainly of articles in the *Journal of Health Economics* and *Health Economics*). The subfield "Economics" accounts for more than 84% of the articles in the discipline "Economics" in our sample. We assigned to "Sociology" (discipline) the subfield "Sociology", to "Psychology" all subfields of the field "Psychology & Cognitive Sciences" as well as the subfield "Psychiatry", to "Other Social Sciences" the field "Social Sciences" (except the subfield "Sociology") as well as all remaining subfields of the field "Economics & Business". Journals in all other fields and subfields were assigned to the residual discipline "Health and other sciences".
We decided not to further subdivide the disciplines in order to have at least 10% of the observations in each category.

[13]Merging was not successful for seven journals (39 articles) in our dataset. For these seven journals, we imputed the discipline based on the first category of Clarivate's Master Journal List. In addition, Science-Metrix does not provide a journal-based classification for 15 multidisciplinary journals in our dataset (65 articles, including e.g. PLOS ONE, Economics & Human Biology, PNAS and Nature). Again, we assign these journals based on the first category of Clarivate's Master Journal List.

et al. 2015). At the lowest level (Level 0), a journal either encourages code sharing or does not specifically mention it. Moving up to Level 1, a *code availability statement* is required in the article, indicating whether the code is accessible and, if so, where it can be found. At Level 2, *mandatory code sharing* is implemented, requiring the code to be posted to a trusted repository, with any exceptions to be identified when the article is submitted. The highest level (Level 3) involves a *reproducibility check*, requiring that the code be posted to a trusted repository and that reported analyses be independently reproduced before publication.

The TOP Factor Database provides the current level of replication code policy for 164 of the 478 journals (34%) in our sample.[14] We searched the websites of the other 315 journals for information on code sharing policies.[15] This search resulted in the *current* state of code sharing policies. However, we are particularly interested in the *historical* development of these policies. We therefore wrote to editors with a code sharing policy level of at least 1 (code availability statement) to ask (i) whether our assessment was correct, and (ii) when this particular policy (and previous code policies) had been adopted. This included the editors of 98 journals in our SOEP-based sample. We received responses from 67 journals (68 %) on the current status of the replication code policy, and 62 journals (63 %) provided information on when the policy was adopted.[16]

2.5. Descriptive statistics

Table 1 provides descriptive statistics to get a better idea of the 2,518 SOEP-based articles in our dataset. Economics is the discipline with the most articles in our sample (45%), followed by sociology (16%), other social sciences (15%), health and other sciences (14%) and psychology (11%). Our dataset includes publications from 1985 to 2021, with an average publication year of 2011, indicating that the majority of articles are from the last decade. In general, the number of SOEP-based publications per year increased strongly over time, from 12 in 1990 (of which 3 in economics) to 26 in 2000 (15 in economics), 118 in 2010 (68 in economics), and 176 in 2020 (55 in economics).

While the average journal impact factor is 1.86 (2.6 for the 5-year impact factor), this is a highly skewed variable with a maximum of 40 (*Nature* in 2016) overall and a maximum

[14] The TOP Factor Database has a community-based approach where registered users provide information about journal policies. Although this approach is not error-free, we found that the information in this database was mostly consistent with the information we obtained from the journals' websites. For all 164 journals, the last update of the code-sharing policy was in 2021 or later.

[15] We focused on *code* sharing policies, and considered *data* sharing policies only when code, syntax files, or computer programs were specifically mentioned as part of "data".

[16] We also double checked the dates provided in Brodeur et al. (2022) and Christensen and Miguel (2018), which focused on *data* availability policies, for 25 and 11 economics journals respectively. In general, we took the information provided by journal websites and editors for granted and did not check the extent to which these policies are actually enforced (see also the discussion in Colliard et al. 2023).

of 9.6 in economics (*Journal of Political Economy* in 2021). Another skewed variable is the number of citations, which ranges from 0 to 4359, with an average of 106. 30 articles have more than 1000 citations, while about 430 articles have less than 10 citations (including 90 with no citations). The vast majority of articles are written in English (87%). The remaining articles are written in German (319 articles; 13%) and French (2 articles; less than 1%). All economics articles in our sample are in English. About 28% are single-author articles, while most articles are written by two authors (38%). About 13% of the articles have four or more authors.

The 2,518 articles are published in 478 different journals (153 in economics, 325 in other disciplines). However, 312 journals (65%) have at most two articles in our dataset, while 14% of journals contribute ten or more articles. The ten most frequent journals in our dataset account for 26% of the articles in our sample. Appendix Table A1 lists the journals with the most articles in our dataset. While previous literature has focused mainly on the top journals in economics, it is evident that we consider a much broader range of journals. While the most common economics journal in our dataset is "Labour Economics" (79 articles), our dataset also includes journals that are ranked higher and journals that are ranked lower. Our dataset also includes 20 articles from TOP-5 economics journals. Similarly, for other disciplines, a wide range of high to low ranked journals is included. Thus, our results do not speak to the situation at the top, but provide a broader view of the state of replication code availability.

3. Results

This section begins by looking at the current state of replication code availability and its correlates in Section 3.1. The following Section 3.2 describes trends in code availability over time and discusses possible explanations for the observed trends. Section **??** examines several factors related to code availability simultaneously, while Section 3.4 explores the relationship between code availability and quality metrics. Finally, Section 3.5 discusses results related to a more rigorous measure of code availability.

3.1. Replication code availability and its correlates

Across all disciplines we found replication code for 151 of the 2,518 SOEP-based publications (6%). Figure 1 shows the availability of code by discipline: We obtained replication code for 43 of the 1,132 SOEP-based articles published in economics journals (3.8 %) and for 108 of the 1,386 articles published in journals of other disciplines (7.8 %). The share of articles with available code is lower in economics than in the other four disciplinary categories in our SOEP-based sample. Psychology has the highest share of articles with available code (almost 10%), followed by other social sciences (9%).

We also find that articles written by a single author are less likely to contain code – and this is true in both economics and other disciplines (see Appendix Figure A1). One possible reason for this observation is that in multi-author collaborations, at least one author is more likely to have a personal website and to place importance on reproducibility. Another explanation could be that single-author articles are less likely to be published in journals with stricter code-sharing policies. A further explanation is that the importance of reproducibility becomes more apparent in multi-author collaborations. In such collaborations, ensuring that all authors obtain identical results by running the code on their respective computers becomes a more straightforward task, underscoring the importance of reproducibility. This would be also in line with the observation of **?** that articles with more authors tend to have better code documentation. Furthermore, code is more likely to be available if the article is written in English (see Figure A2 in the Appendix).[17] Again, stricter code-sharing policies in English-speaking journals could be an explanation for this observation. Section 3.3 examines whether these bivariate correlations also hold when controlling for other factors, including publication year dummies and indicators for the journal's code sharing policy.

Next, we look more closely at the articles that provide code: What software is used, how are others informed of the availability of the code, where is the code stored? The majority of articles with code provided use Stata (87%), a proprietary software, followed by R with 21% (see Appendix Figure A3).[18] SPSS, MATLAB, MPLUS and SAS each account for less than 4% of studies with code. These proportions are roughly comparable to the figures from the 2018 SOEP user survey, which asked respondents about the statistical packages they use to analyse SOEP data: About 77% use Stata, 24 % use R, 23 % use SPSS, 11 % use other (Britzke and Schupp 2019).[19] Similarly, in a collection of more than 8000 economic articles with replication packages, (Kranz 2023) finds that Stata is the most commonly used software package (71.6%), followed by Matlab (24.5%) and R (9.8%).

In both economics and other disciplines, most studies that provide code use the journal website to provide information about code availability (see Figure 2). However, for a sizable proportion of studies, code availability was only disclosed via the authors' websites (32 % in economics and 18 % in other disciplines), highlighting the importance of also checking authors' websites for code availability: Looking only at journal websites would lead to a substantial underestimation of code availability. In general, only a few studies disseminate

[17]It is interesting to see, that all articles in economics journals are published in English.

[18]In economics, code is only provided in Stata (95%), Matlab (9%) and R (7%) in our sample. The percentages add up to more than 100 because some articles rely on more than one software program.

[19]The shares are similar in the 2017 SOEP user survey with 76% Stata, 31 % R, 19 % SPSS, 10 % other (Britzke and Schupp 2019).

the code via online repositories exclusively, without referring to the availability of the code via the journal or the authors' websites (1 article in economics, 4 articles in other disciplines). It is also interesting to see that it seems less common in economics to disclose the availability of the code both through the journal and the authors' websites (5% vs. 33%). Economics also differs from other disciplines in terms of where the code provided is actually stored (see Figure A4 in the Appendix). While for the majority of studies in economics code can be downloaded directly from author or journal websites, in other disciplines 50% of studies with code rely on the repository of the Open Science Framework (OSF), compared to less than 5% in economics.[20]

3.2. Developments in code availability over time

Figure 3 shows the development of replication code availability across publication years, starting with the first SOEP-based peer-reviewed publications in 1985 (1989 in economics) until 2021. The figure shows several striking patterns. First, by 2012 code is available for only five studies in economics (0.85 %) and three (0.55 %) in other disciplines. In economics, the first publication with available code is from 2007 (Alesina and Fuchs-Schündeln 2007), while overall the earliest code we obtained is from 1995 (Rendtel et al. 1995). Here the SPSS code was printed as an appendix to the paper itself, highlighting the challenges researchers faced in sharing code before the widespread use of the World Wide Web, which allowed code to be stored on journal or private websites or in dedicated online repositories. The first digital code we found was from Kohler (1998). Second, the figure reveals that from 2012 onward the annual share of publications with code increased strongly for both economics and other disciplines. The increase from less than 2 % in 2012 to 17 % and 18 % in 2021 in economics and other disciplines, respectively, suggests a huge change in code-sharing behavior. In economics, the biggest growth happened after 2019, while in the other fields the strongest increase occurred a few years earlier. Economics was lagging behind, but is catching up in recent years.

When the non-economics category is broken down by discipline in Figure A5 in the Appendix, it can be seen that the share of publications with code was similar across disciplines until around 2015. After 2015, however, the increase in sociology, psychology and other social sciences was earlier and stronger than in economics.

Next, we look at how the mode of code availability disclosure and the location of storage have changed over time. The proportion of studies for which we found information on code

[20]The OSF is a network for researchers to document their research projects, e.g. by registering analysis plans and sharing their material. It also allows a Digital Object Identifier (DOI) to be assigned to the code provided, making it possible to make code permanently available.

availability through the journal increased strongly over the years, both in economics and in other disciplines (see Appendix Figure A6). There is also an increase in the proportion of studies for which we found information on code availability only through the authors' websites. However, this increase is less pronounced, suggesting that journals play a major role in the increase in code availability over time. While in other disciplines there is an increase in the disclosure of code availability through both the journal and the authors' websites, there is no such development in economics. Appendix Figure A7 highlights the sharp increase in the use of OSF from 2015 onwards in other disciplines, but not in economics.[21]

There are several possible explanations for the remarkable increase in code availability over time. First, technological advances have reduced the financial and logistical barriers to code sharing. These developments include the widespread use of individual websites by researchers, the establishment of journal websites, and the creation of specialized online repositories.[22] Prior to these technological advances, code could only be stored and shared on floppy disks or by printing codes in appendices (as in Rendtel et al. 1995). The increased use of repositories to store code (Figure A7) and the substantial proportion of studies for which code is available on the authors' websites (Figure 2) speak to the importance of technological advances in increasing code sharing.

Second, evolving journal policies on code sharing may play an important role in explaining the positive trend in code availability over time. To explore this explanation in more detail, we collected information on journal policies on code sharing. Appendix Figure A8 shows the distribution of current (as of November 2023) code policies for the journals (upper panels) and the article (lower panels) in our sample. In both economics and other disciplines, about 80% of journals have a Level-0 policy, meaning that the journal either encourages code sharing or does not mention code sharing at all. While Level-1 policies (stating whether code is available and, if so, where) are slightly less common in economics than in other disciplines, in both groups about 10% of journals require that code be deposited in a trusted repository (Level 2). However, the highest (and strictest) level is more common in economics journals. At this level, not only must the code be deposited in a trusted repository, but all the results of the article must be independently reproduced.

We then examine how these code-sharing policies have evolved over time for the 478 journals in our sample (see the upper panels of Figure 1). The *American Economic Review*

[21] In line with these findings, OSF users and usage have increased exponentially since its inception in 2012 (see Nosek et al. 2022).

[22] Miguel (2021: 196) highlights the importance of these online repositories, saying that they "have been so successful that it is easy to forget what an important innovation the professional curation, storage, and management of research data and code has been."

was the first economics journal in our dataset to announce an explicit code-sharing policy (Bernanke 2004).[23] In economics, the share of journals requiring mandatory code sharing has increased continuously since 2004, with a particularly strong increase in the period 2015-2019. By 2021, more than 20% of the economics journals in our sample (31 out of 153) have a mandatory code-sharing policy. In other disciplines, there is also a notable increase in Level-2 policies over time, which kicked in somewhat later.[24] By 2021, 10% of non-economics journals require the provision of replication code. Similarly, we see that at the end of our observation period, the share of journals with reproducibility checks (Level 3) is higher in economics.[25]

While the upper panels of Figure 1 look at the journal level, the lower panels of this figure examine how many of the articles in our dataset were published when a code-sharing policy was in place. The first such articles are, in our sample, from 2007 in economics and from 2014 in other disciplines. These two dates roughly coincide with the first available replication codes (ignoring the two articles with code from the 1990s), speaking to the relevance of these policies. After these dates, the share of articles that were covered by a mandatory code-sharing policy increased steadily, up to more than 11% in economics and to more than 8% in other disciplines in 2021. There are only three articles in our sample that were published when a reproducibility check was in place (Level 3), all of them in economics.

The importance of journal policies for code provision is underlined by the fact that four of the five economics articles published in or before 2012 that provide code (see Figure 3) are published in journals with mandatory code sharing policies (three in the *American Economic Review* and one in the *Review of Economic Studies*).[26]

Apart from technological a and journal policies, a third explanation for the increase in code availability over time another contributing factor could be the heightened awareness

[23]The *Journal of Political Economy* announced a similar Level-2 policy in 2005 and the *Review of Economic Studies* in 2006. This was followed by the *Brookings Papers on Economic Activity* in 2007, the *Canadian Journal of Economics* in 2008, and the *American Economic Journals* in 2009.

[24]While the upper panels of Figure 4 show that some economic journals are rather front-runners in terms of code sharing policies, the first journal with an explicit code sharing policy was the *Stata Journal*. The journal, due to its focus on software code, requires since the first issue in 2001 that software code is made available through the journal's website.

[25]The *Biometrical Journal* was, in 2009, the first journal in our sample to implement a policy of independently reproducing the results of accepted manuscripts prior to publication (Hofner et al. 2016). However, code sharing is not mandatory: "the Editor in Chief notifies the authors ... already during the review process of the journal's RR [reproducible research] policy and strongly encourages the submission of code and data to make the article reproducible" (Hofner et al. 2016: 418).

[26]The fifth article is published in the *Journal of Applied Econometrics*, which also strongly emphasizes the replicability of results by other researchers. While the journal has a mandatory data sharing policy (if the data are not confidential), code sharing is only encouraged but not mandatory according to the journal website. Therefore, their code sharing policy is classified as Level 0.

among researchers, driven by, for instance, negative examples of retracted studies and the discussion about a reproducibility and replicability crisis. Romero (2019) traces the beginning of the discussion around the replicability crisis to several events in the early 2010s. This broadly coincides with the increased availability of replication code in our sample. Moreover, starting in the 2010s several comprehensive studies documented that many published articles cannot be replicated; this includes psychology (Open Science Collaboration 2015), cancer research (Nosek and Errington 2017), experiments in economics (Camerer et al. 2016) and social sciences in general (Camerer et al. 2018). This increased awareness of the replicability crisis may have contributed both to more authors sharing their code, making their studies easier to reproduce, and to more journals implementing code (and data) sharing policies. In economics, the Herndon et al. (2013)-Reinhart and Rogoff (2010) debate has also contributed to a growing awareness of reproducibility and code availability.

While Figure 4 provides evidence that journal policies have contributed to this development, it is important to note that less than 5% of articles in both economics and other disciplines were published when a Level 1 or higher code-sharing policy was in place. Furthermore, Figure 5 shows that there is also an increase in code availability over time net of journal policies. This figure excludes publications that were published when the journal had a code-sharing policy of Level 1 or higher. We take this as indirect evidence that bottom-up initiatives by individual researchers to make their code available, even when not required by the journal, have also contributed to the increase in code availability.

3.3. Regression results: Predictors of code availability

In this section, we use regression analysis to examine several factors related to code availability simultaneously. More specifically, we estimate linear probability models (LPMs) with code availability as the binary dependent variable, in which we sequentially include the previously discussed factors.[27] Table 2 shows the results. The first column of Table 2 is based on the bivariate associations from Figure 1 between code availability and disciplines. Economics is the reference category and it can be seen that the coefficients for all four

[27] More specifically, we estimate equations of the following form

$$Y_p = \alpha + X_p'\beta + \sum_{t=1986}^{2021} \delta_t \cdot \mathbb{1}(year = t) + \varepsilon_p, \tag{1}$$

where Y_p is a binary variable indicating whether the code is publicly available for publication p. X_p is the vector of explanatory variables (e.g., indicators for the journal's discipline); δ_t is a set of dummy variables for each publication year. Finally, ε_p is the error term. We use heteroscedasticity-robust standard errors throughout our analysis, as the variance of ε_p may not be constant across different values of the explanatory variables.

other disciplines are positive, indicating that among SOEP-based articles code provision is less common in economics. The coefficients for all four disciplines are reduced when time dummies are included in column (2). This suggests that the lower prevalence of code-sharing in economics is partly due to the higher proportion of economics articles in publications from earlier years, when code-sharing was generally less prevalent. While all four discipline coefficients are smaller in column (2), the coefficients are still jointly statistically significant.

This picture remains the same when including indicator variables for whether the article is in English and whether the article was written by a single author (column 3). While the bivariate associations showed that code provision is generally higher among articles in English language (see Figure A2), this association disappears when publication year dummies are included. This indicates that the lower share of articles in other languages with provided code is due to the fact that non-English articles were mostly published in earlier years when code provision was less common. The "single-author"-coefficient is negative and statistically significant at the 10% level: Articles written by a single author are less likely to provide code, even after controlling for discipline, year of publication, and language of the article.

Column (4) considers journal policies on code sharing, with a Level 0 policy as the reference category. The coefficients for Level 1 (Code Availability Statement), Level 2 (Mandatory Code Sharing), and Level 3 (Reproducibility Check) are all positive, indicating that they are positively related to code sharing compared to a policy that says nothing about code sharing or only encourages it. The coefficients for Level 2 and Level 3 are particularly large (also relative to all other coefficients in the table), suggesting that a mandatory code-sharing policy (which is also part of Level 3) greatly increases the availability of code. The coefficient for Level 2 suggests that a mandatory code-sharing policy is associated with an increase in code availability of about 46 percentage points.[28] While this is a considerable magnitude, it is also not close to 100. This is consistent with previous findings in the literature that mandatory code policies are not always strictly enforced (Colliard et al. 2023; Vlaeminck and Herrmann 2015). It is interesting to see that the coefficients for the other disciplines increase when controlling for journal policies on code sharing. This indicates that code sharing is lower in economics despite the stricter code sharing policies.

3.4. Code availability and quality metrics

In this section, we examine the relationship between code availability and quality metrics. These quality metrics include both the impact factor of the journal in which an article was

[28]While this coefficient is statistically significant at common levels, the coefficient for Level 3 is not statistically significant despite its size. This is likely due to the small number of articles in our sample that are subject to a Level 3 policy.

published and citation measures of the article. The left panel of Figure 6 shows that the average journal impact factor for SOEP-based publications in economics is about 2, when the article provide code, while it is about 1.3 when no code is provided. The picture is very similar when looking at all other disciplines (right panel of Figure 6) or when looking at 5-year impact factors (see Appendix Figure A9): Code-sharing is correlated with higher impact factors. Multiple regression analyses, presented in Appendix Table A2, confirm this finding. The regression results show that, code availability is associated with a 0.95 point higher JIF in economics and a 0.51 point higher JIF in other disciplines, controlling for year fixed effects. This relationship does not change much when controlling for the single-author indicator, the language of the article, and the journal's code-sharing policy. Code sharing is correlated with higher impact factors - and we cannot explain this correlation with the included predictor variables.

Figure 6 shows the average number of citations in Google Scholar for studies with and without available code.[29] It shows that in both economics and other disciplines, articles with replication code available receive, on average, more citations than articles without code. This relationship is more pronounced in economics than in other disciplines. Table A3 in the Appendix shows average marginal effects from related Poisson regressions. The table confirms that articles with available code receive more citations, also when controlling for the single author indicator and the language of the article. When controlling for the journal policies and the journal impact factor, the coefficients drop in size but are still meaningful: Articles with available code receive, ceteris paribus, on average 71 more citations in economics and about 49 more in other disciplines.

As with the impact factor, we cannot completely explain the positive correlation between code provision and citations with the included predictor variables. However, the conditional associations presented in this section should not be interpreted as causal effects of code provision, as we do not control for many potentially confounding factors (e.g. author characteristics and article quality).[30]

3.5. Robustness

While in the previous sections we focused on whether any replication code was available, we did not consider the quality of the code provided in terms of its ability to reproduce the results. The quality of code is important as several studies have shown that code availability does not ensure computational reproducibility (see e.g. Chang and Li 2017; Eubank 2016;

[29]Studies published in early years have more time to accumulate citations. Therefore, Figure 6 controls for year of publication dummies.

[30]Christensen et al. (2019) provide some evidence that *data* sharing increases citations when using changes in data sharing policies as instrumental variables.

Fišar et al. ming; Gertler et al. 2018; McCullough et al. 2006), mostly due to missing data, non-transferability of code from one computer to the replicator's computer, unstable results after repeated code runs, and differences between published and reproduced results (including copying errors). However, it is beyond the scope of this paper to analyze the reproducibility of all 151 SOEP papers with code provided. For example, Gertler et al. (2018) report that in their attempt to reproduce the results of published papers in economics, they set a limit of four hours for each paper to try to run the code (excluding code runtime).

As a first step in looking at the quality of the code provided, we work with a stricter definition of code availability. More specifically, we count a publication as providing code only if the code loads raw SOEP data, which means that replication packages that do not start with the raw SOEP data (as provided by the SOEP group) but with some intermediate "analysis" dataset are not counted. It is likely that such an intermediate dataset will not allow the exact reproduction of the final results, since many hidden researcher decisions relate to sample restriction and variable definition choices.

We find that about a quarter of the publications that provide some code do not start from the raw SOEP data, but from some intermediate dataset. This means that, using the stricter measure of replication code availability, only 4.5% of the publications in our dataset provide code, or 113 in absolute terms. Figure 8 shows the evolution of code availability (according to the stricter definition) over time. Comparing this figure to Figure 3, which is based on our main definition of code availability, several observations stand out. First, the overall pattern is the same: code availability increases over time in both economics and other disciplines. Second, the first study with available code that loads the SOEP raw data is from 2007 (12 years later than the first publication with any replication code), and the first publication outside of economics that provides code according to the stricter definition is from 2012 (Kohler et al. 2012).

While none of the studies providing code met the stricter definition in the period before 2007, the proportion of studies with code meeting the stricter definition remained relatively constant over time (75 % in 2007-2011, 76.7 % in 2012-2016, 75.7 % in 2017-2021). This suggests that code quality, as measured by our crude measure, has not improved much over time. However, it also suggests that the increased rate of code provision over time did not come at the expense of the quality of the code provided. The Online Appendix **??** repeats all our analyses, this time using the stricter definition of code availability. Generally, the patterns are very similar to those with our main definition of code availability.

4. Conclusion

In his famous article "Let's take the con out of econometrics" Edward Leamer (1983) heavily criticized the state of empirical economic research, culminating in the quote "hardly anyone takes anyone else's data analysis seriously". If nobody takes data analysis of other researcher serious, it is a serious problem for a profession that strongly relies on empirical results and calls for evidence-based policy-making. However, Angrist and Pischke (2010) argue that the Leamer critique is no longer justified in empirical microeconomics due to the commitment to rigorous research designs for identifying causal effects. This is what they label the credibility revolution in their aptly titled article "The credibility revolution in empirical economics: How better research design is taking the con out of econometrics". However, recent studies by Huntington-Klein et al. (2021) and Breznau et al. (2022) highlight a disconcerting phenomenon: researchers drawing disparate conclusions when analyzing the same data to answer identical questions, even when they apply rigorous research designs for the identification of causal effects. This is due to the many decisions that researchers must make during data collection, preparation, and analysis, also known as the researcher's degree of freedom. It suggests that relying solely on credible research designs does not fully address the lingering concerns voiced by Leamer (1983). Making data and code used in the empirical analysis publicly available is one way to make the researcher's degree of freedom more transparent and to examine how crucial it is for a given research question. Having access to data and code would allow others to check all the details of the empirical analyses and run additional sensitivity checks.

Against this backdrop, this study looks at how often researchers share the computer code they use for their studies, focusing on publications that use the German Socio-Economic Panel (SOEP) dataset—a widely-used data source in economics and other disciplines. We searched extensively to find out if the code was available for all 2,518 articles using SOEP in peer-reviewed journals. We checked the journal websites, the articles themselves, the websites of the authors, and specific online repositories.

The analysis shows that only 6% of SOEP studies have accessible replication code, a low proportion, particularly considering that code provision is crucial for the reproducibility of scientific results and associated with little costs for the researchers. However, there is a positive trend as the share of studies with available code increased strongly over time in both economics and other disciplines. We argue that the increase in code provision over time is driven by a mixture of three factors: Technological advances (including websites of authors and journals as well as the establishment of specific repositories), top-down initiatives of journals, e.g. mandating code sharing, and bottom-up initiatives of individual researchers, who post code on their individual websites. Additionally, we find that studies with accessible code tend to be published in journals with higher impact factors and receive more citations, emphasizing the potential of code sharing as a quality signal in academic research.

Our analysis can serve as a first step for further analyses. We sketch three of such further analyses. First, we examined whether studies provide replication code, but we did not study whether the code actually allowed the results to be reproduced. We take a first step in this direction, by applying the stricter definition of code availability where the code must load raw SOEP data. However, it is beyond the scope of this paper to analyze the reproducibility of all SOEP papers with code provided. Second, another potential avenue for future research is to go more in the direction of causality. While we document correlations between journal policies and code availability, future research could evaluate the impact of code sharing policies on code availability with more rigorous methods for causal inference. Similarly, future studies could look at the effects of code provision on citation metrics.

Third, future studies could also examine code availability beyond the SOEP. A possible approach would be to investigate the development of code availability in articles published in leading journals across various disciplines. However, this sampling method may not offer a comprehensive overview of code sharing as it concentrates solely on the most prestigious journals. Further, in many cases the provision of code will not be enough to allow for reproducibility as data are missing and there are substantial differences between disciplines regarding the proportion of data that are shareable with the research community. While publications based on the SOEP are not a random sample of studies in economics or other disciplines, a complimentary strategy would be to draw random samples from publications in different disciplines. However, this procedure presents several challenges. Firstly, it is necessary to verify whether the study is an empirical one and the rates of empirical studies differ between disciplines. Secondly, the disciplines vary in terms of the proportion of publications based on primary or secondary data analyses. Thirdly, it is unclear what the population is and how to construct a sampling frame. By focusing on the SOEP, we avoid many of these issues as the SOEP levels the playing field as the data set is publicly available by researchers around the world. In this setting code availability is the most crucial ingredient for reproducibility.

We conclude this paper by arguing for what we call the Credibility Revolution 2.0. This new paradigm contends that the mere focus on research designs is not enough to address the criticisms of Leamer (1983). Instead, it calls for a fundamental shift toward code and data sharing as the new normal. In our view, the true test of credibility lies not only in the design of the research, but in the ability of peers to scrutinize every detail, facilitating collective validation of results through reproducibility and transparency. Providing replication code is a low-cost measure in this new paradigm, and basically any empirical researcher can engage in it, paving the way for a more trustworthy empirical economics.

Acknowledgments:
We thank Toman Barsbai, Markus Jäntti, and Michael Kvasnicka for helpful comments and feedback. We are also indebted to Niema El Moufti, Ashwin Fernandez, and Alexander Grawe for outstanding research assistance. We are grateful to Uta Rahmann of the SOEP group for providing us with the SOEPlit database and for answering our questions about SOEPlit. A big thank you also goes to all the journal editors who took the time to answer our questions about the implementation of specific code sharing policies.

Data and Code Availability Statement:
It would be very strange if you researched code availability but did not make your code available. Therefore: The code is available at XXX. We will provide the data and a more detailed replication package once our manuscript is accepted for publication.

References

Alesina, A. and N. Fuchs-Schündeln (2007). Good-bye Lenin (or not?): The effect of communism on people's preferences. *American Economic Review 97*(4), 1507–1528.

Angrist, J. D. and J.-S. Pischke (2010). The credibility revolution in empirical economics: How better research design is taking the con out of econometrics. *Journal of Economic Perspectives 24*(2), 3–30.

Archambault, É., O. H. Beauchesne, and J. Caruso (2011). Towards a multilingual, comprehensive and open scientific journal ontology. In *Proceedings of the 13th international conference of the international society for scientometrics and informetrics*, Volume 13, pp. 66–77. Durban, South Africa.

Bernanke, B. S. (2004, March). Editorial statement. *American Economic Review 94*(1), 404.

Blanco-Perez, C. and A. Brodeur (2020). Publication bias and editorial statement on negative findings. *The Economic Journal 130*(629), 1226–1247.

Breznau, N., E. M. Rinke, A. Wuttke, H. H. Nguyen, M. Adem, J. Adriaans, A. Alvarez-Benjumea, H. K. Andersen, D. Auer, F. Azevedo, et al. (2022). Observing many researchers using the same data and hypothesis reveals a hidden universe of uncertainty. *Proceedings of the National Academy of Sciences 119*(44), e2203150119.

Britzke, J. and J. Schupp (Eds.) (2019). *SOEP Wave Report 2018.* Number 2018 in SOEP Wave Report. Berlin.

Brodeur, A., N. Cook, and C. Neisser (2022). P-hacking, data type and data-sharing policy. *IZA Discussion Paper 15586.*

Camerer, C. F., A. Dreber, E. Forsell, T.-H. Ho, J. Huber, M. Johannesson, M. Kirchler, J. Almenberg, A. Altmejd, T. Chan, E. Heikensten, F. Holzmeister, T. Imai, S. Isaksson, G. Nave, T. Pfeiffer, M. Razen, and H. Wu (2016). Evaluating replicability of laboratory experiments in economics. *Science 351*(6280), 1433–1436.

Camerer, C. F., A. Dreber, F. Holzmeister, T.-H. Ho, J. Huber, M. Johannesson, M. Kirchler, G. Nave, B. A. Nosek, T. Pfeiffer, et al. (2018). Evaluating the replicability of social science experiments in Nature and Science between 2010 and 2015. *Nature Human Behaviour 2*(9), 637–644.

Chang, A. C. and P. Li (2017). A preanalysis plan to replicate sixty economics research papers that worked half of the time. *American Economic Review, Papers & Proceedings 107*(5), 60–64.

Cholewiak, S. A., P. Ipeirotis, V. Silva, and A. Kannawadi (2021). Scholarly: Simple access to google scholar authors and citation using python.

Christensen, G., A. Dafoe, E. Miguel, D. A. Moore, and A. K. Rose (2019). A study of the impact of data sharing on article citations using journal policies as a natural experiment. *PloS one 14*(12), e0225883.

Christensen, G. and E. Miguel (2018). Transparency, reproducibility, and the credibility of economics research. *Journal of Economic Literature 56*(3), 920–980.

Colliard, J.-E., C. Hurlin, and C. Pérignon (2023). The economics of research reproducibility. *HEC Paris Research Paper FIN-2019-1345 (December 20, 2023).*

Donohue, John J., I. and S. D. Levitt (2001, 05). The impact of legalized abortion on crime. *The Quarterly Journal of Economics 116*(2), 379–420.

Eubank, N. (2016). Lessons from a decade of replications at the Quarterly Journal of Political Science. *PS: Political Science and Politics 49*(2), 273–276.

Fišar, M., B. Greiner, C. Huber, E. Katok, A. I. Ozkes, and Management Science Reproducibility Collaboration (forthcoming). Reproducibility in Management Science. *Management Science.*

Foote, C. L. and C. F. Goetz (2008). The impact of legalized abortion on crime: Comment. *The Quarterly Journal of Economics 123*(1), 407–423.

Gertler, P., S. Galiani, and M. Romero (2018). How to make replication the norm. *Nature 554*(7693), 417–419.

Goebel, J., M. M. Grabka, S. Liebig, M. Kroh, D. Richter, C. Schröder, and J. Schupp (2019). The German Socio-Economic Panel (SOEP). *Jahrbücher für Nationalökonomie und Statistik 239*(2), 345–360.

Hardwicke, T. E. and J. P. Ioannidis (2018). Mapping the universe of registered reports. *Nature Human Behaviour 2*(11), 793–796.

Herndon, T., M. Ash, and R. Pollin (2013, 12). Does high public debt consistently stifle economic growth? A critique of Reinhart and Rogoff. *Cambridge Journal of Economics 38*(2), 257–279.

Hofner, B., M. Schmid, and L. Edler (2016). Reproducible research in statistics: A review and guidelines for the biometrical journal. *Biometrical Journal 58*(2), 416–427.

Huntington-Klein, N., A. Arenas, E. Beam, M. Bertoni, J. R. Bloem, P. Burli, N. Chen, P. Grieco, G. Ekpe, T. Pugatch, M. Saavedra, and Y. Stopnitzky (2021). The influence of hidden researcher decisions in applied microeconomics. *Economic Inquiry 59*(3), 944–960.

Key, E. M. (2016). How are we doing? data access and replication in political science. *PS: Political Science & Politics 49*(2), 268–272.

Kohler, U. (1998). Zur Attraktivität der Grünen bei älteren Wählern. *Kölner Zeitschrift für Soziologie und Sozialpsychologie 50*, 536–559.

Kohler, U., M. Ehlert, B. Grell, J. P. Heisig, A. Radenacker, and M. Wörz (2012). Verarmungsrisiken nach kritischen Lebensereignissen in Deutschland und den USA. *KZfSS Kölner Zeitschrift für Soziologie und Sozialpsychologie 64*(2), 223–245.

Kranz, S. (2023, dec). Usage shares of programming languages in economics research.

Laurinavichyute, A., H. Yadav, and S. Vasishth (2022). Share the code, not just the data: A case study of the reproducibility of articles published in the journal of memory and language under the open data policy. *Journal of Memory and Language 125*, 104332.

Leamer, E. E. (1983). Let's take the con out of econometrics. *The American Economic Review 73*(1), 31–43.

McCullough, B. D., K. A. McGeary, and T. D. Harrison (2006). Lessons from the JMCB Archive. *Journal of Money, Credit and Banking 38*(4), 1093–1107.

Miguel, E. (2021). Evidence on research transparency in economics. *Journal of Economic Perspectives 35*(3), 193–214.

Nosek, B. A., G. Alter, G. C. Banks, D. Borsboom, S. D. Bowman, S. J. Breckler, S. Buck, C. D. Chambers, G. Chin, G. Christensen, M. Contestabile, A. Dafoe, E. Eich, J. Freese, R. Glennerster, D. Goroff, D. P. Green, B. Hesse, M. Humphreys, J. Ishiyama, D. Karlan, A. Kraut, A. Lupia, P. Mabry, T. Madon, N. Malhotra, E. Mayo-Wilson, M. McNutt, E. Miguel, E. L. Paluck, U. Simonsohn, C. Soderberg, B. A. Spellman, J. Turitto, G. VandenBos, S. Vazire, E. J. Wagenmakers, R. Wilson, and T. Yarkoni (2015). Promoting an open research culture. *Science 348*(6242), 1422–1425.

Nosek, B. A. and T. M. Errington (2017). Making sense of replications. *Elife 6*, e23383.

Nosek, B. A., T. E. Hardwicke, H. Moshontz, A. Allard, K. S. Corker, A. Dreber, F. Fidler, J. Hilgard, M. Kline Struhl, M. B. Nuijten, J. M. Rohrer, F. Romero, A. M. Scheel, L. D. Scherer, F. D. Schönbrodt, and S. Vazire (2022). Replicability, robustness, and reproducibility in psychological science. *Annual Review of Psychology 73*(1), 719–748.

Open Science Collaboration (2015). Estimating the reproducibility of psychological science. *Science 349*(6251), aac4716.

Reinhart, C. M. and K. S. Rogoff (2010). Growth in a time of debt. *American Economic Review 100*(2), 573–78.

Rendtel, U., G. G. Wagner, and J. R. Frick (1995). Eine Strategie zur Kontrolle von Längsschnittgewichtungen in Panelerhebungen: das Beispiel des Sozio-Oekonomischen Panels (SOEP). *Allgemeines Statistisches Archiv 79*(3), 252–277.

Romero, F. (2019). Philosophy of science and the replicability crisis. *Philosophy Compass 14*(11), e12633.

SOEP Group (2024). Data access.

Vlaeminck, S. and L.-K. Herrmann (2015). Data policies and data archives: A new paradigm for academic publishing in economic sciences? In B. Schmidt and M. Dobreva (Eds.), *New Avenues for Electronic Publishing in the Age of Infinite Collections and Citizen Science. Proceedings of the 19th International Conference on Electronic Publishing*, Amsterdam, pp. 145–155. IOS Press.

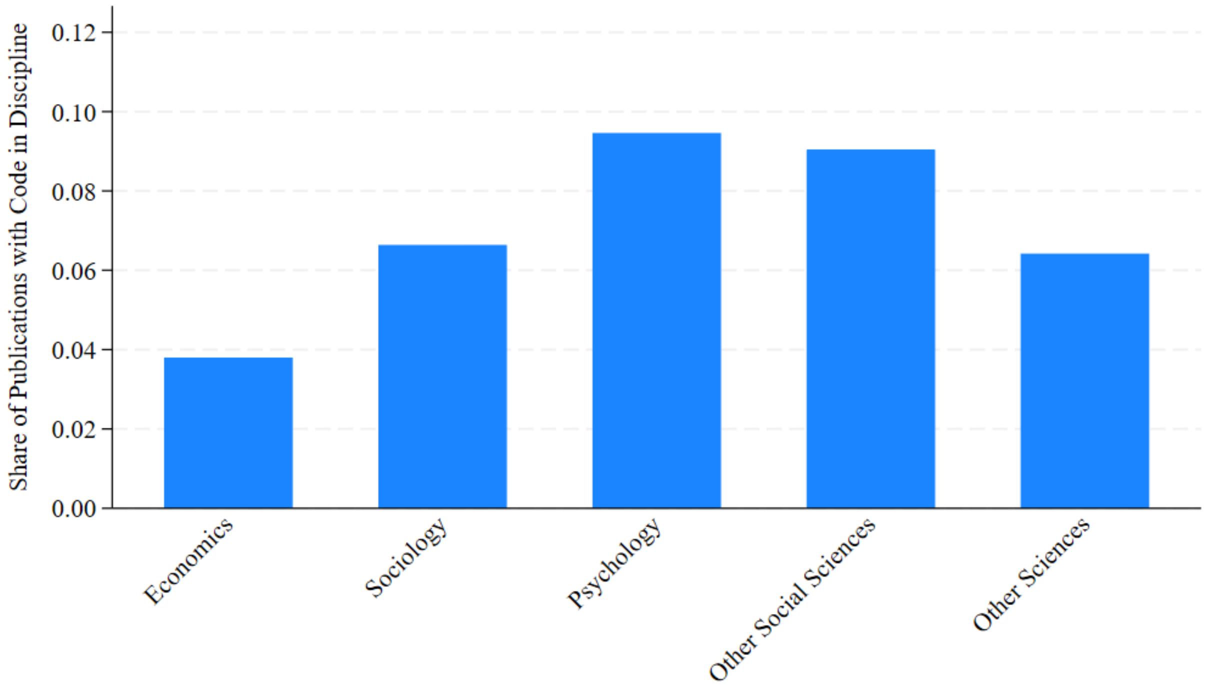
5. Figures and Tables

Table 1: Descriptive statistics

	(1) Mean	SD	Min	Max
Economics	0.45	0.50	0.00	1.00
Sociology	0.16	0.36	0.00	1.00
Psychology	0.11	0.31	0.00	1.00
Other Social Sciences	0.15	0.36	0.00	1.00
Health and Other Sciences	0.14	0.34	0.00	1.00
Publication Year	2011.91	7.38	1985.00	2021.00
Journal Impact Factor (JIF)	1.86	2.20	0.04	40.14
5 Year JIF	2.60	2.58	0.09	43.77
Citation Count from Google Scholar	106.19	242.25	0.00	4359.00
Language of Article is English	0.87	0.33	0.00	1.00
Single Author	0.28	0.45	0.00	1.00
Two Authors	0.38	0.49	0.00	1.00
Three Authors	0.21	0.41	0.00	1.00
Four or More Authors	0.13	0.33	0.00	1.00

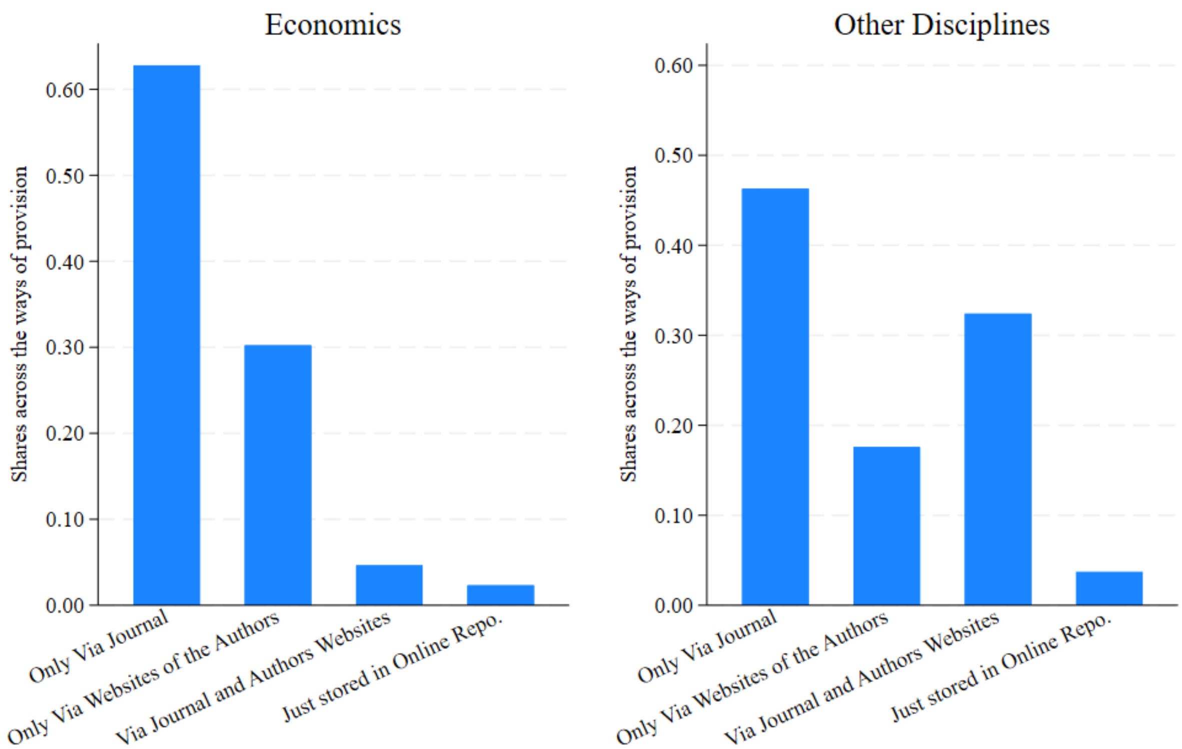
Note: The table displays mean, standard deviation, minimum and maximum for different variables in our sample of 2,518 SOEP-based publications.
Source: Own calculations.

Figure 1: Publications with publicly available code by discipline



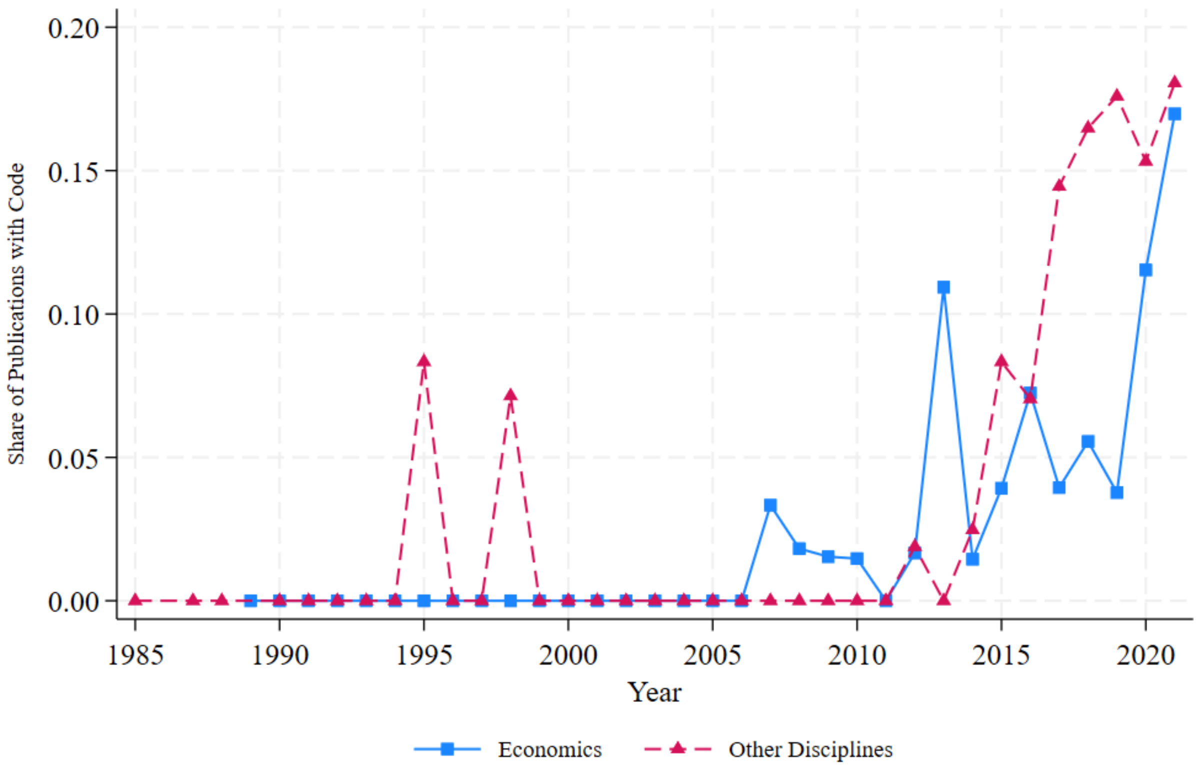
Note: This figure shows the share of SOEP-based publications with publicly available replication code by discipline. This figure and all other figures consider only peer-reviewed publications in journals that are listed in the citation indices of Clarivate Analytics.
Source: Own calculations, based on SOEPlit and own research on code availability.

Figure 2: Publications with publicly available code by mode of disclosure



Note: This figure shows the share of SOEP-based publications with publicly available replication code by the mode how others are informed about the availability of code. The left panel relates to the situation in economics, while the right panel relates to all other disciplines.
Source: Own calculations, based on SOEPlit and own research on code availability.

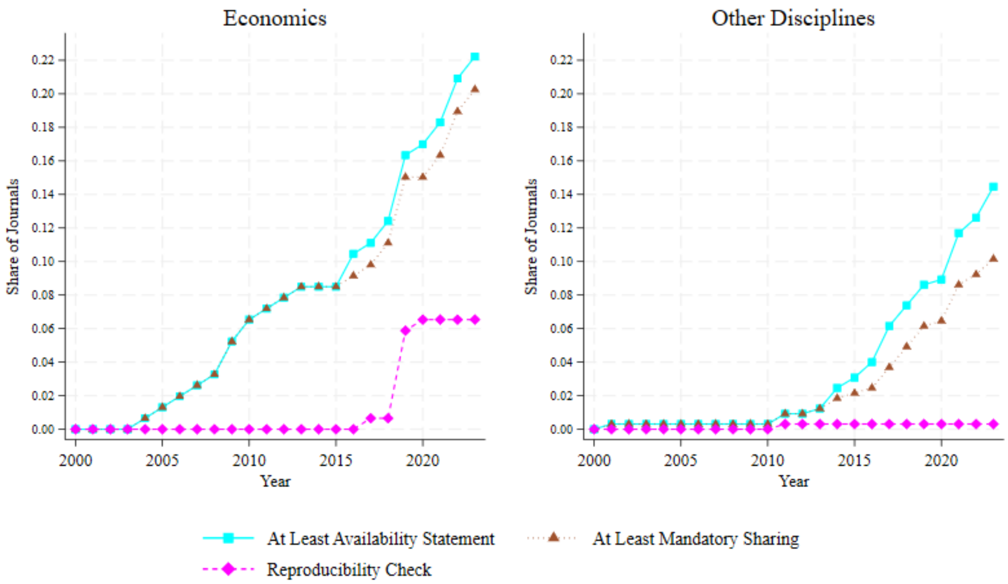
Figure 3: Publications with publicly available code over time



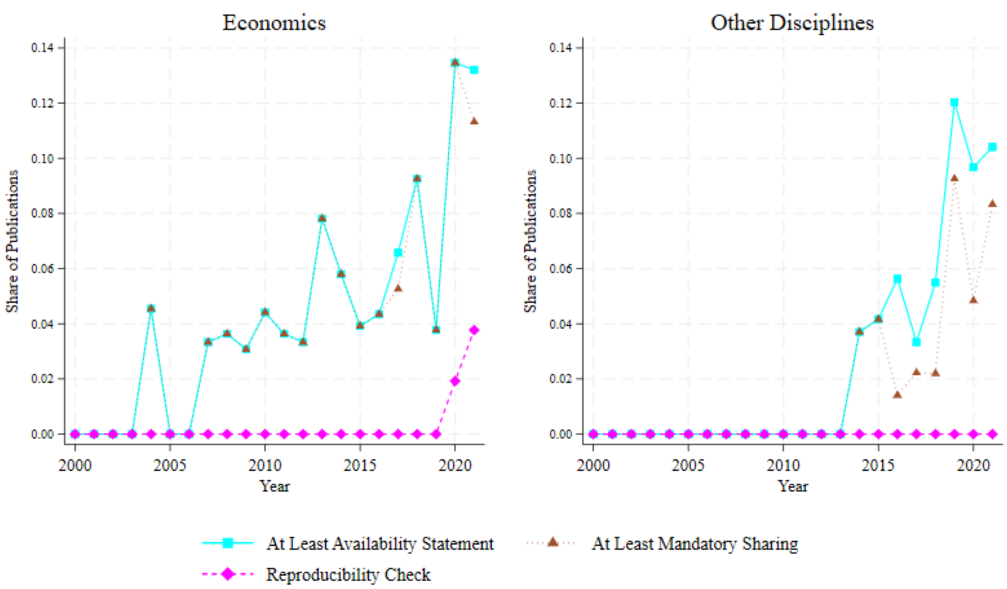
Note: This figure shows the share of SOEP-based publications with publicly available replication code across publication years. The figure considers only peer-reviewed publications in journals that are listed in the citation indices of Clarivate Analytics.
Source: Own calculations, based on SOEPlit and own research on code availability.

Figure 4: Development of code sharing policies over time

(a) Journal level

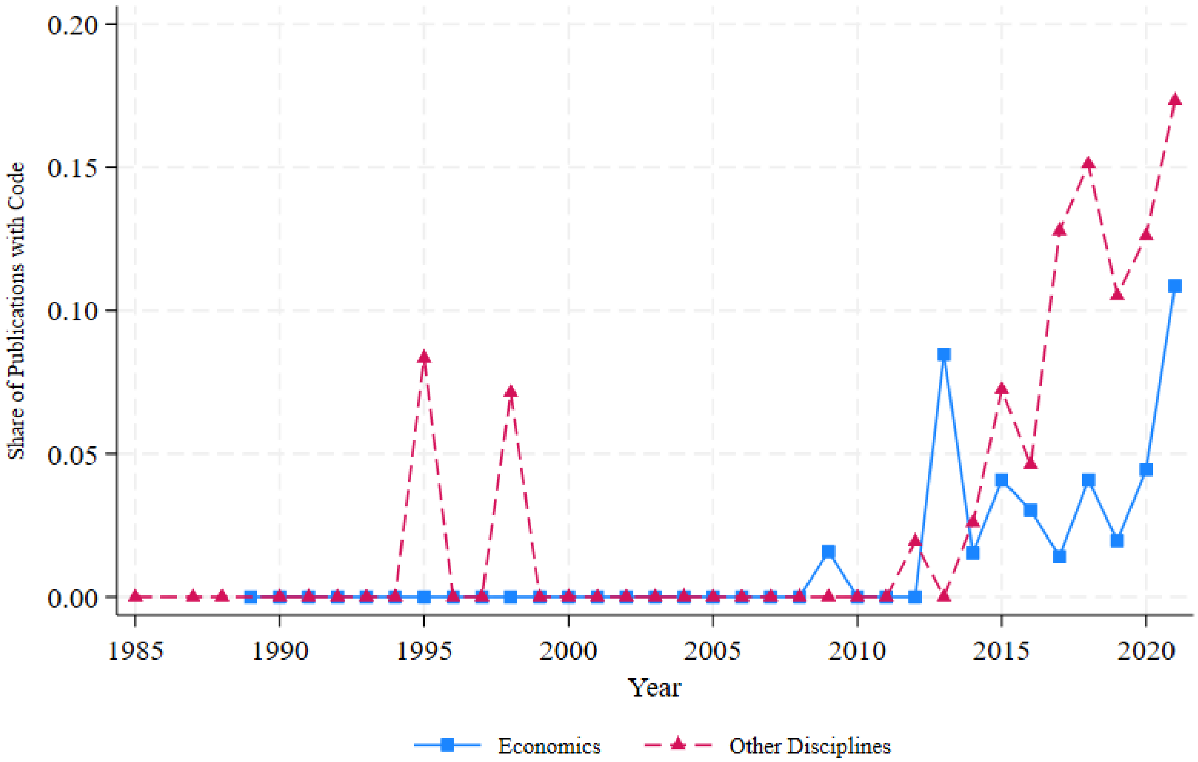


(b) Article level



Note: This figure shows the development of replication code policies over time for the journal (upper panels) and articles (lower panels) in our sample. The left panels relate to the situation in economics, while the right panels relate to all other disciplines.
Source: Own calculations, based on SOEPlit, TOP Factor Database, and own research on journals' replication code policies.

Figure 5: Publicly available code, net of journal policies



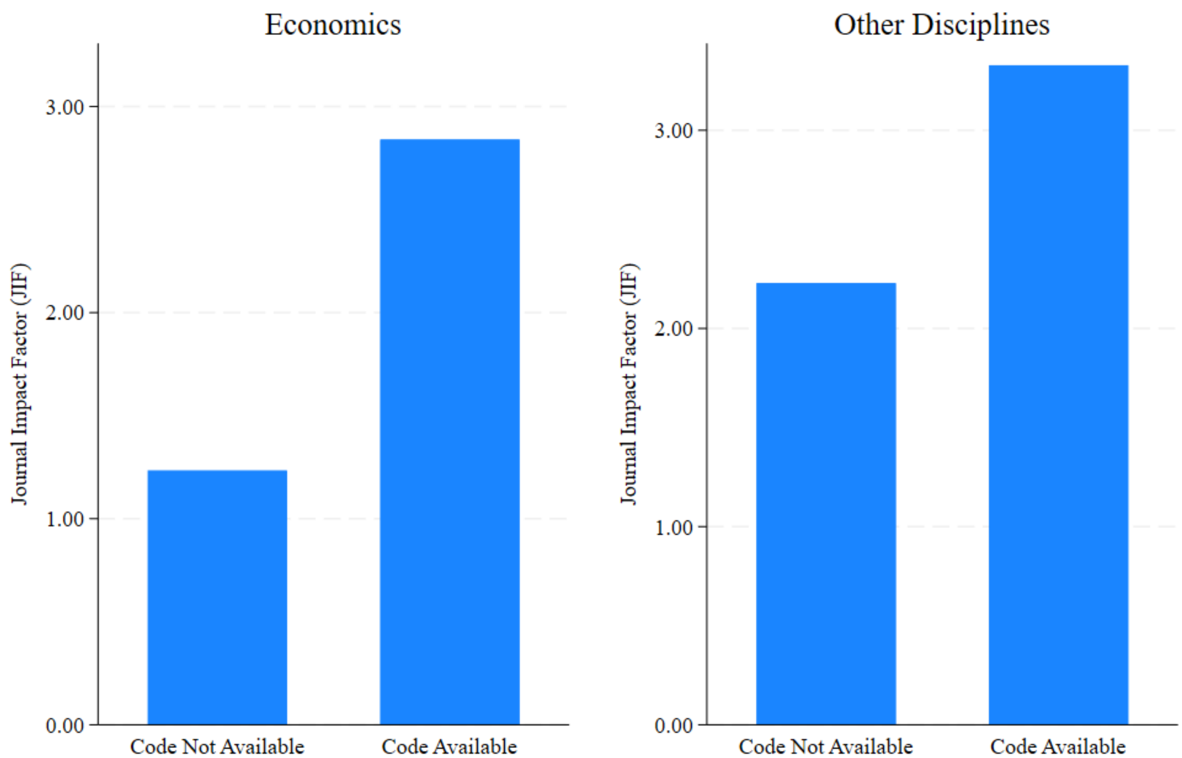
Note: This figure shows the development of the share of SOEP-based publications with publicly available replication code, excluding publications that were published when the journal had a code-sharing policy in place.
Source: Own calculations, based on SOEPlit and own research on code availability.

Table 2: Predictors of code availability

	Dependent Variable: Code Available (Yes/No)				
	(1)	(2)	(3)	(4)	(5)
Sociology	0.028**	0.025*	0.027*	0.044***	0.046***
	(0.014)	(0.014)	(0.015)	(0.015)	(0.015)
Psychology	0.057***	0.027	0.026	0.029*	0.005
	(0.019)	(0.018)	(0.018)	(0.017)	(0.017)
Other Social Sciences	0.052***	0.029*	0.030*	0.042***	0.040***
	(0.016)	(0.015)	(0.016)	(0.014)	(0.015)
Other Sciences	0.026*	0.000	-0.001	-0.002	0.004
	(0.014)	(0.015)	(0.015)	(0.015)	(0.014)
English Article			0.000	-0.001	-0.002
			(0.014)	(0.014)	(0.014)
Single Author			-0.018*	-0.010	-0.012
			(0.009)	(0.009)	(0.009)
Code Availability Statement				0.045	
				(0.076)	
Mandatory Code Sharing				0.367***	
				(0.053)	
Reproducibility Check				0.558**	
				(0.279)	
Publication Year FE	No	Yes	Yes	Yes	Yes
Only Level-0 Journals	No	No	No	No	Yes
F (4, 2517)	5.18***	1.96***	1.96***	4.64***	3.92***
Adj. R^2	0.01	0.05	0.05	0.14	0.04
N	2518	2518	2518	2500	2057

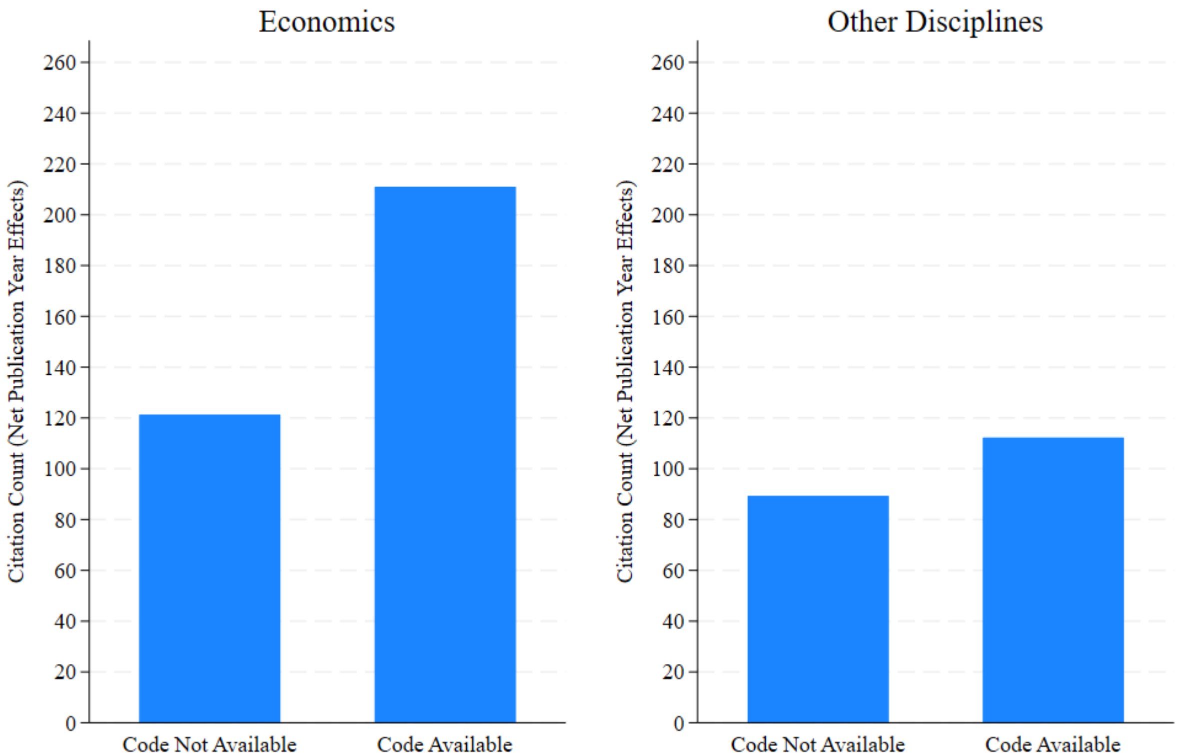
Note: The table shows coefficients for linear probability models with code availability as binary dependent variable and robust standard errors in parentheses. The base categories for the discipline dummies is economics. $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$..
Source: Own calculations, based on SOEPlit as well as own research on code availability and journals' replication code policies.

Figure 6: Code availability and journal impact factor



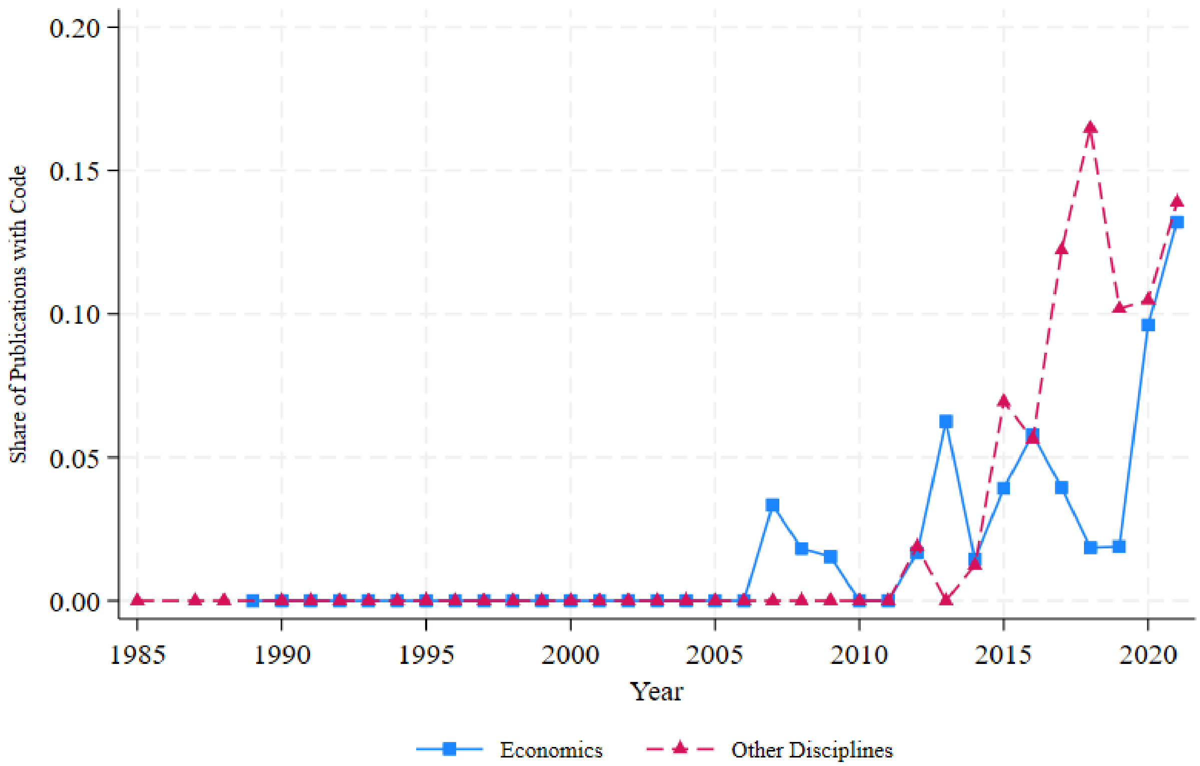
Note: This figure shows the average journal impact factor separately for SOEP-publications with and without provided code. The left panel relates to all publications in economics, while the right panel relates to publications in all other disciplines.
Source: Own calculations, based on SOEPlit, journal impact factors from Clarivate Analytics, and own research on code availability.

Figure 7: Code availability and citations



Note: This figure shows the average citations in Google Scholar for SOEP-publications with and without provided code. The left panel relates to all publications in economics, while the right panel relates to publications in all other disciplines.
Source: Own calculations, based on SOEPlit, Google Scholar, and own research on code availability.

Figure 8: Publicly available code that loads raw SOEP data over time



Note: This figure shows the development of the share of SOEP-based publications with publicly available replication code that loads raw SOEP data. The figure considers only peer-reviewed publications in journals that are listed in the citation indices of Clarivate Analytics.
Source: Own calculations, based on SOEPlit and own research on code availability.

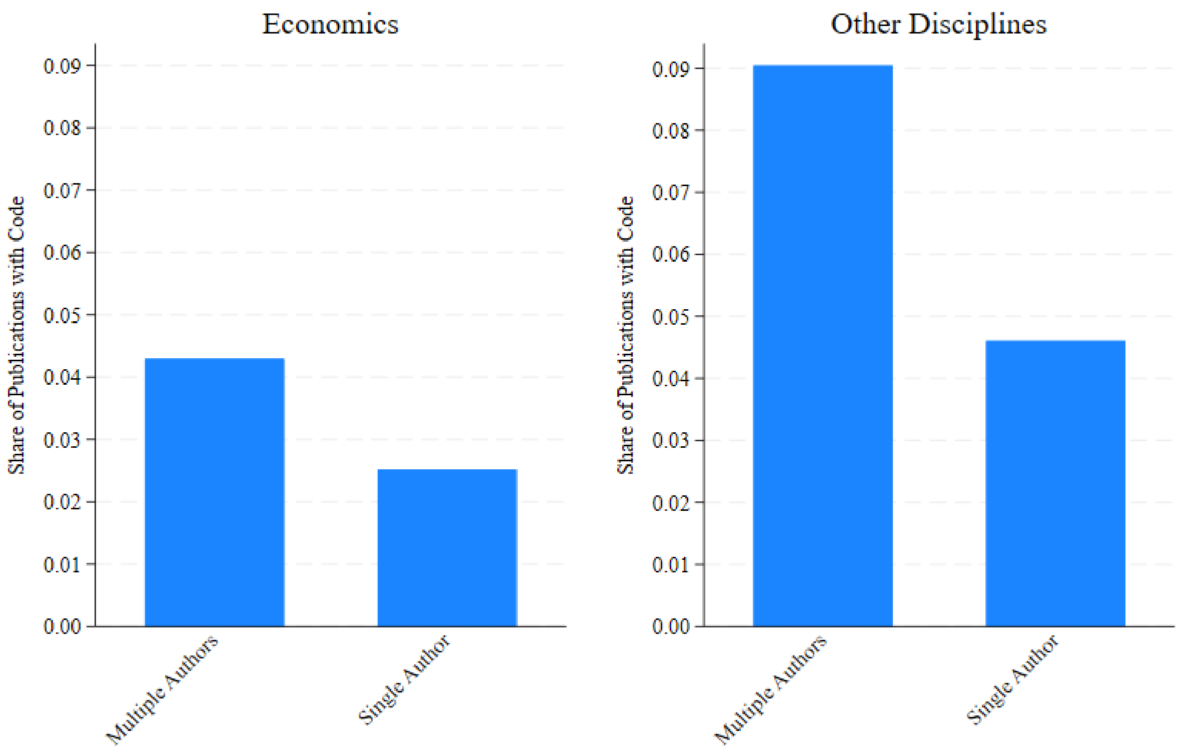
A. Appendix

Table A1: Most frequent journals in our sample

Economics	
Labour Economics	79
Jahrbücher für Nationalökonomie und Statistik	62
Review of Income and Wealth	49
Economics Letters	48
Journal of Population Economics	44
Journal of Economic Behavior & Organisation	43
Health Economics	30
German Economic Review	29
Applied Economics	26
Empirical Economics	25
Economic Journal	24
Journal of Health Economics	22
European Economic Review	21
Journal of Human Resources	21
Sociology	
Kölner Zeitschrift fur Soziologie und Sozialpsychologie	98
European Sociological Review	94
Zeitschrift für Soziologie	67
Soziale Welt	17
Research in Social Stratification and Mobility	16
Psychology	
Social Indicators Research	76
Journal of Happiness Studies	31
Journal of Personality and Social Psychology	20
Psychology and Aging	20
Journal of Research in Personality	14
Other social sciences	
Journal of Marriage and Family	29
Zeitschrift für Familienforschung	26
Journal of European Social Policy	18
Small Business Economics	18
Demography	17
Health & other sciences	
AStA-Advances in Statistical Analysis	35
PLOS ONE	24
Gesundheitswesen	19
Social Science & Medicine	19
Journals of Gerontology Series B-Psychological Sciences	13

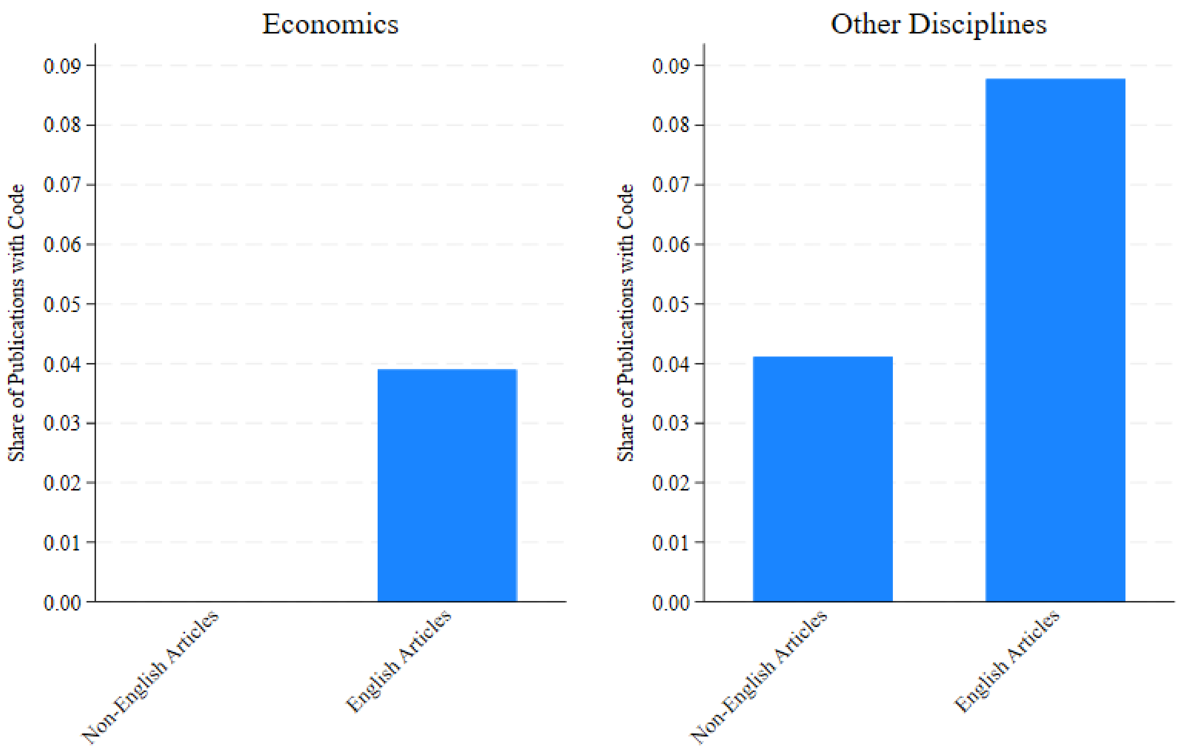
Note: The table lists, for each of our five disciplinary categories, the names of the journals with the most articles in our dataset and the corresponding number of articles.
Source: Own counting, based on SOEPlit.

Figure A1: Code availability: Single author vs. multiple authors



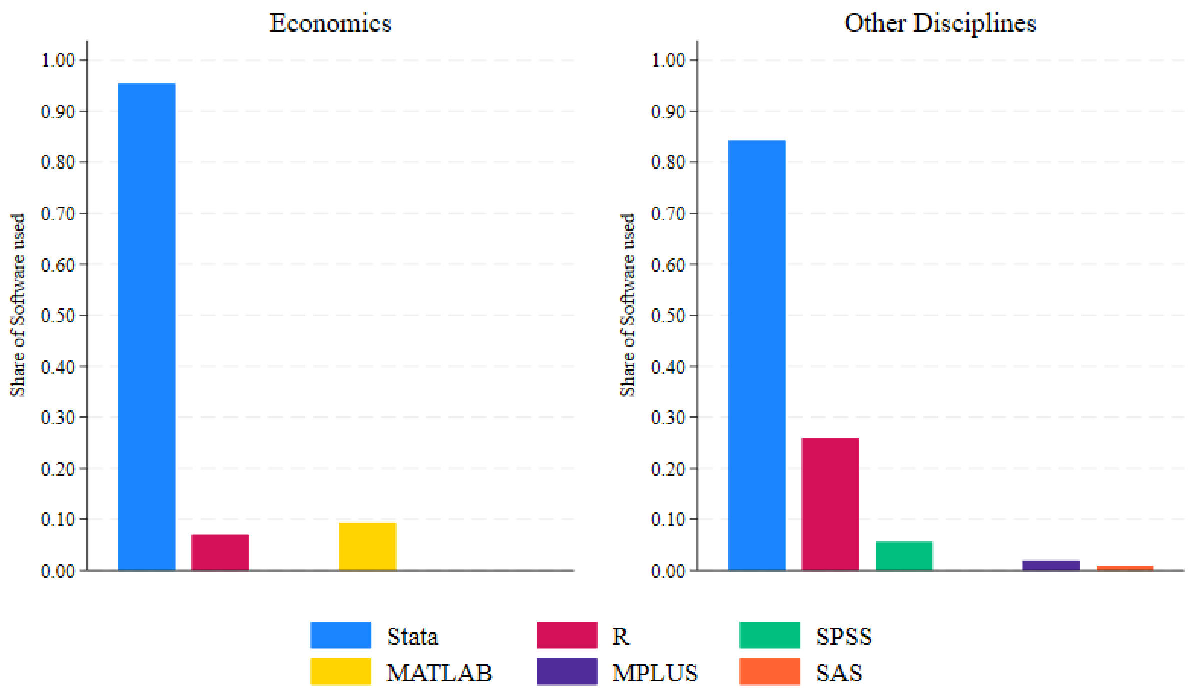
Note: This figure shows the share of SOEP-based publications with publicly available replication code, separately for articles written by a single author and articles with multiple authors. The left panel relates to economics, while the right panel relates to all other disciplines.
Source: Own calculations, based on SOEPlit and own research on code availability.

Figure A2: Code availability: Language of the article



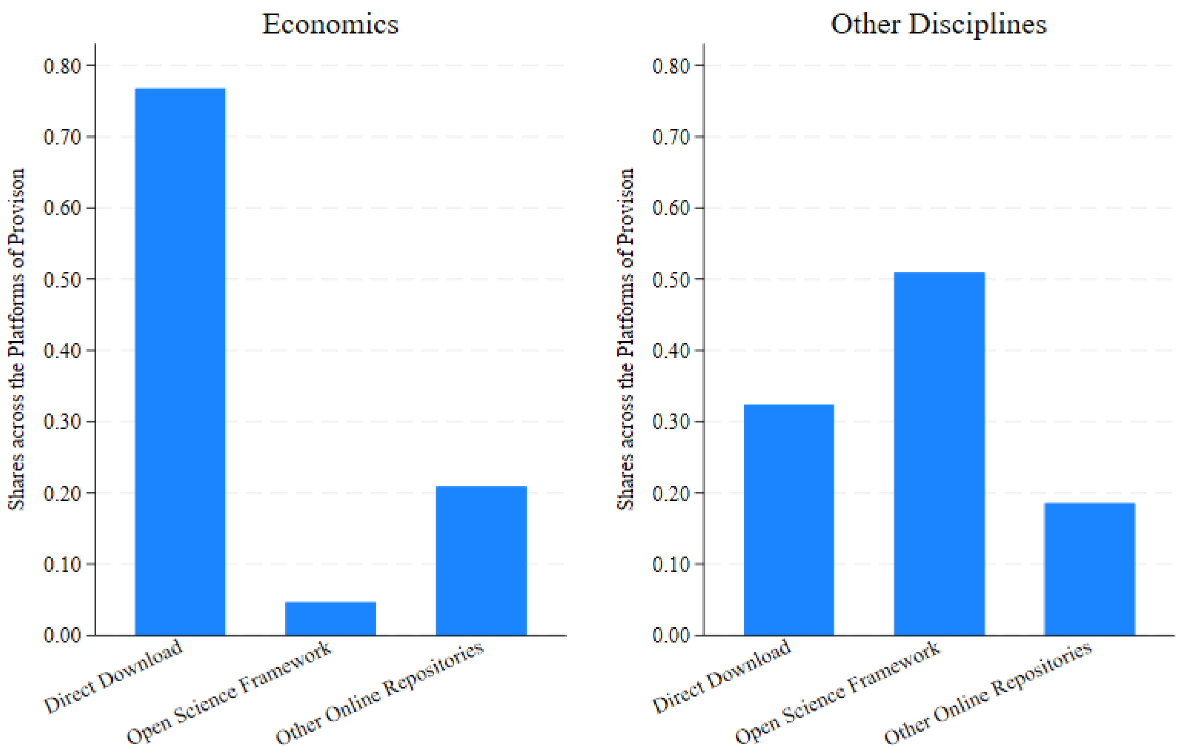
Note: This figure shows the share of SOEP-based publications with publicly available replication code, separately for articles in English and in all other languages (mostly German). The left panel relates to economics, while the right panel relates to all other disciplines.
Source: Own calculations, based on SOEPlit and own research on code availability.

Figure A3: Software used in the replication packages



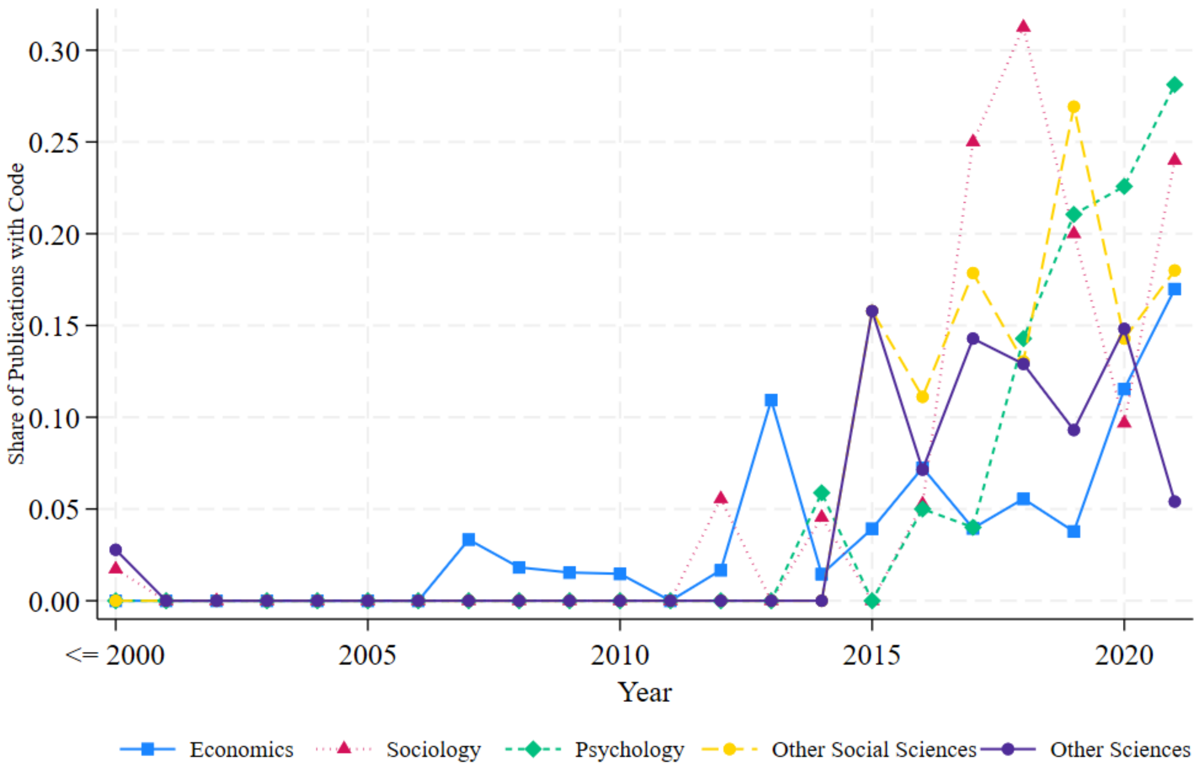
Note: This figure shows the shares across the software used in the available code packages in our sample. The left panel relates to economics, while the right panel relates to all other disciplines. Note that the shares in each panel do not sum up to 1 as sometimes more than one statistical software is used.
Source: Own calculations, based on SOEPlit and own research on code availability.

Figure A4: Publications with publicly available code by storage location



Note: This figure shows the share of SOEP-based publications with publicly available replication code by the location in which the code is stored. The left panel relates to economics, while the right panel relates to all other disciplines.
Source: Own calculations, based on SOEPlit and own research on code availability.

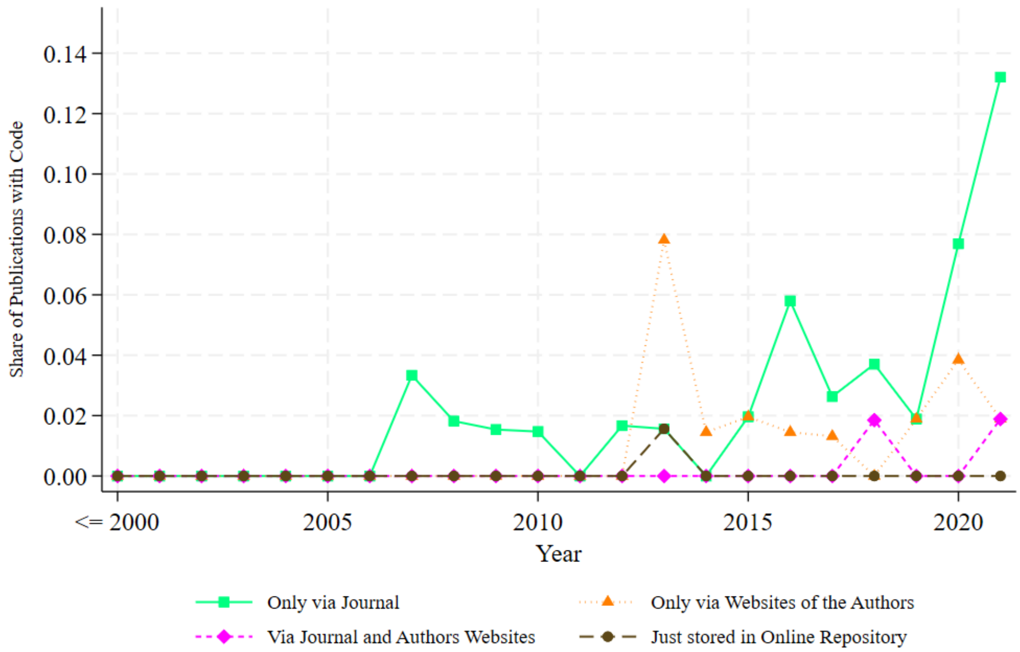
Figure A5: Share of SOEP-based publications with publicly available code, by discipline



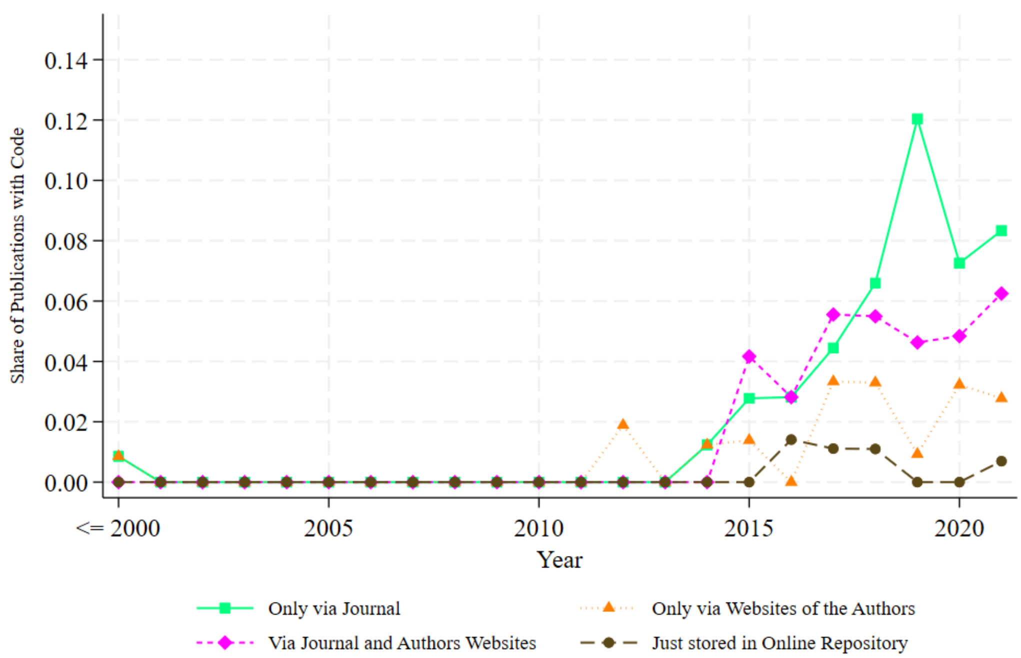
Note: This figure shows the share of SOEP-based publications with publicly available replication code across publication years and for various disciplines. The figure considers only peer-reviewed publications in journals that are listed in the citation indices of Clarivate Analytics. The left endpoints of the graph represent the average over all years prior to 2001, as the availability of code before this year is marginal.
Source: Own calculations, based on SOEPlit and own research on code availability.

Figure A6: Mode of code availability disclosure over time

(a) Economics



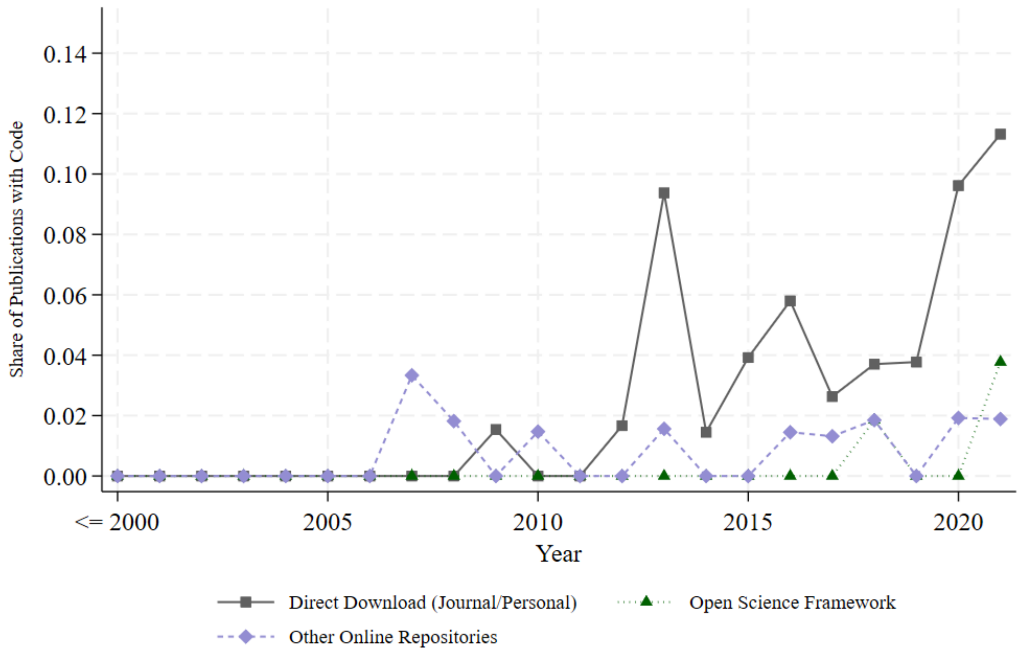
(b) Other disciplines



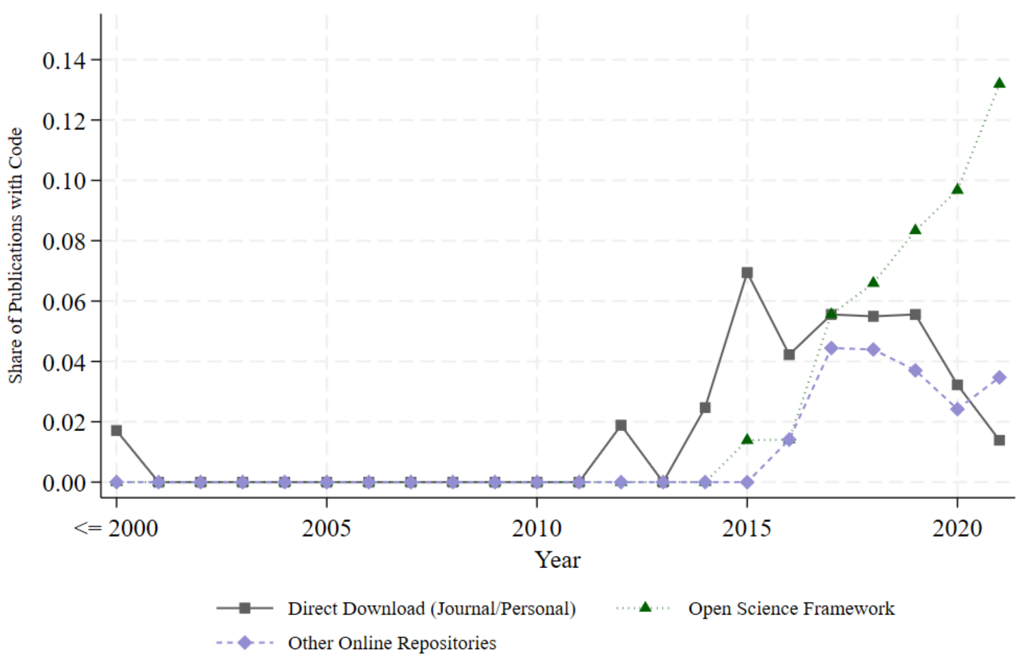
Note: This figure shows the share of SOEP-based publications with publicly available replication code across publication years, differentiating whether the code was provided on the journal webpages or another way (i.e., author webpage or repository) or both ways. Panel a) refers to economics, while panel b) refers to all other disciplines in our sample. The figure considers only peer-reviewed publications in journals that are listed in the citation indices of Clarivate Analytics. The left endpoints of the graph represent the average over all years prior to 2001, as the availability of code before this year is marginal.
Source: Own calculations, based on SOEPlit and own research on code availability.

Figure A7: Publications with publicly available code by storage location, over time

(a) Economics

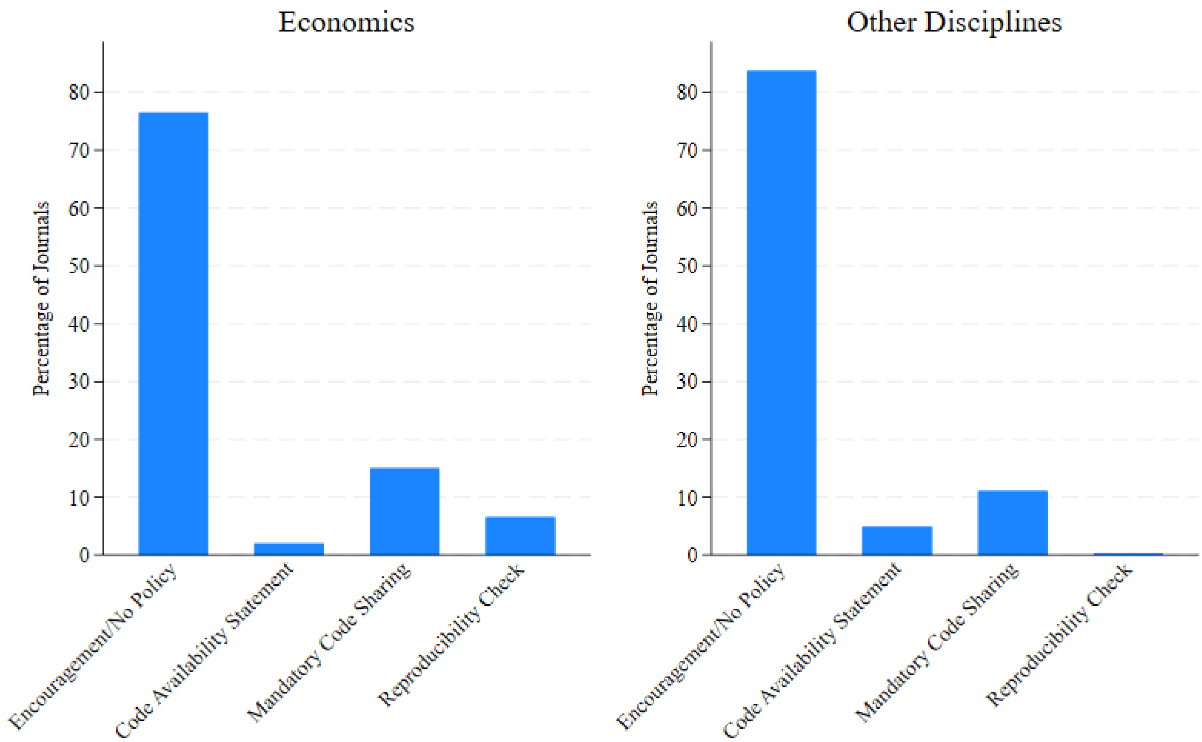


(b) Other disciplines



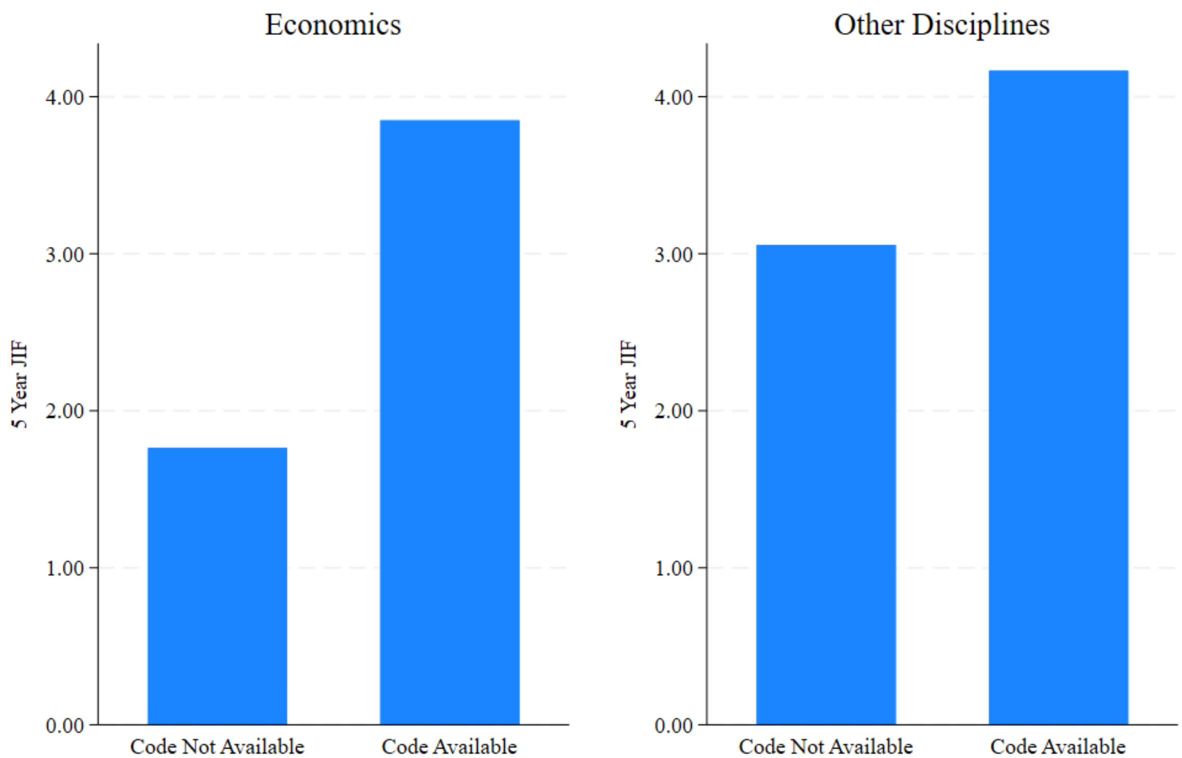
Note: This figure shows the share of SOEP-based publications with publicly available replication code across publication years, differentiating whether the code was stored directly on journal/personal webpages, in the OSF repository, or in another repository. Panel a) refers to economics, while panel b) refers to all other disciplines in our sample. The figure considers only peer-reviewed publications in journals that are listed in the citation indices of Clarivate Analytics. The left endpoints of the graph represent the average over all years prior to 2001, as the availability of code before this year is marginal.
Source: Own calculations, based on SOEPlit and own research on code availability.

Figure A8: Current state of code sharing policies



Note: This figure shows the distribution of replication code policies for the 478 journals in our sample. The left panel relates to the situation in economics, while the right panel relates to all other disciplines.
Source: Own calculations, based on SOEPlit, TOP Factor Database, and own research on journals' replication code policies.

Figure A9: Code availability and the 5-year journal impact factor



Note: This figure shows the average 5-year journal impact factor separately for SOEP-publications with and without provided code. The left panel relates to economics, while the right panel relates to all other disciplines. *Source*: Own calculations, based on SOEPlit, journal impact factors from Clarivate Analytics, and own research on code availability.

Table A2: Code availability and the 2-year journal impact factor

	Dependent Variable: Journal Impact Factor (JIF)				
	(1)	(2)	(3)	(4)	(5)
Code Available	1.46***	0.64***			
	(0.19)	(0.18)			
Code X Economics			0.95***	0.98***	0.61**
			(0.31)	(0.31)	(0.30)
Code X Other Discipline			0.51**	0.48**	0.35
			(0.20)	(0.20)	(0.22)
Sociology			-0.01	0.49***	0.53***
			(0.06)	(0.08)	(0.08)
Psychology			1.09***	1.13***	1.02***
			(0.09)	(0.09)	(0.11)
Other Social Sciences			0.32***	0.42***	0.46***
			(0.09)	(0.08)	(0.08)
Other Sciences			1.60***	1.76***	1.67***
			(0.24)	(0.24)	(0.21)
English Article				1.08***	1.07***
				(0.10)	(0.09)
Single Author				-0.24***	-0.20***
				(0.05)	(0.05)
Code Availability Statement					4.97**
					(2.32)
Mandatory Code Sharing					0.72***
					(0.21)
Reproducibility Check					2.71**
					(1.16)
Publication Year FE	No	Yes	Yes	Yes	Yes
Adj. R^2	0.03	0.17	0.24	0.27	0.31
N	2426	2426	2426	2426	2408

Note: The table presents estimates from OLS regressions with the 2-year journal impact factor as the dependent variable and robust standard errors in parentheses. Column (1) confirms the pattern from Figure 6: Code provision is associated with higher impact factors and this association remains, albeit to a lesser extent, when controlling for publication year dummies (column 2). In column (3), we additionally control for the discipline of the journal and replace the code availability indicator with separate code availability indicators for economics and for other disciplines. The picture remains similar. Controlling for year fixed effects, code availability is associated with a 0.95 point higher JIF in economics and a 0.51 point higher JIF in other disciplines. These relationships do not change much when controlling for the single-author indicator and the language of the article (column 4). The relationship also remains very similar when controlling for the journal's code-sharing policy (column 5).
$^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$.. The base category for the discipline dummies is economics.
Source: Own calculations, based on SOEPlit, journal impact factors from Clarivate Analytics, as well as own research on code availability and journals' replication code policies.

Table A3: Code availability and Google Scholar Citations

	Dependent Variable: Citation Count					
	(1)	(2)	(3)	(4)	(5)	(6)
Code Available	-30.3*	96.3**				
	(15.7)	(39.2)				
Code X Economics			159.4**	173.1**	41.9	26.9
			(77.7)	(79.2)	(52.2)	(45.8)
Code X Other Discipline			55.5**	56.8**	24.2	29.1
			(24.6)	(23.0)	(24.7)	(24.6)
Sociology			-54.7***	-3.3	1.2	-11.2
			(9.6)	(11.1)	(11.1)	(12.0)
Psychology			45.1**	42.7**	46.3**	24.9
			(22.7)	(20.4)	(20.5)	(17.5)
Other Social Sciences			-34.9***	-26.4***	-22.6**	-32.3***
			(10.5)	(9.8)	(9.7)	(10.9)
Other Sciences			-55.3***	-39.5***	-36.7***	-73.9***
			(10.5)	(10.2)	(9.8)	(11.6)
English Article				84.8***	85.2***	78.9***
				(8.0)	(8.0)	(8.7)
Single Author				-34.9***	-33.0***	-29.5***
				(8.4)	(8.4)	(8.4)
Code Availability Statement					124.5	-88.5***
					(146.0)	(23.9)
Mandatory Code Sharing					175.7*	155.4*
					(94.0)	(87.9)
Reproducibility Check					592.0	384.1*
					(361.9)	(206.5)
Journal Impact Factor						15.4***
						(2.5)
Publication Year FE	No	Yes	Yes	Yes	Yes	Yes
Pseudo R^2	0.00	0.25	0.29	0.34	0.36	0.42
N	2518	2518	2518	2518	2500	2408

Note: The table presents average marginal effects from Poisson regressions with robust standard errors in parentheses and citations as the dependent variable. While the raw relationship between code availability and citations is negative (column 1), the coefficient reverses sign when publication year fixed effects are included (column 2), highlighting the importance of controlling for differences in publication years. Column (3) introduces controls for different disciplines and replaces the code availability indicator with separate code availability indicators for economics and for other disciplines. Column (4) additionally controls for the single author indicator, while Column (5) considers the journal policies on code sharing as well. Column (6) controls for the JIF.
$^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$.. The base category for the discipline dummies is economics.
Source: Own calculations, based on SOEPlit, Google Scholar, as well as own research on code availability and journals' replication code policies.