# Geographical Bias in GitHub: Perceptions and Reality

Ayushi Rastogi
IIIT-Delhi
Delhi, India
ayushir@iiitd.ac.in

Nachiappan Nagappan
Microsoft Research
Redmond, USA
nachin@microsoft.com

Georgios Gousios
Radboud University Nijmegen
The Netherlands
g.gousios@cs.ru.nl

## ABSTRACT

Open source development has often been considered to be a level playing field for all developers. But there has been little work to investigate if bias plays any role in getting contributions accepted. The work presented in this study tries to understand the influence of geographical location on the evaluation of pull requests in GitHub - one of the primary open source development platforms.

Using a mixed-methods approach that analyzes 70,000+ pull requests and 2,500+ survey responses, we find that geographical location explains statistically significant differences in pull request acceptance decisions. Compared to the United States, submitters from United Kingdom (22%), Canada (25%), Japan (40%), Netherlands (43%), and Switzerland (58%) have higher chances of getting their pull requests accepted. However, submitters from Germany (15%), Brazil (17%), China (24%), and Italy (19%) have lower chances of getting their pull requests accepted compared to the United States. The probability of pull request acceptance decisions increase by 19% when the submitter and integrator are from the same geographical location. Survey responses from submitters indicate the perception of bias is strong in Brazil and Italy matching our results. Also, 8 out of every 10 integrators feel that it is easy to work with submitters from the same geographical location.

## Keywords

Empirical software engineering; software process; empirical studies

## Categories and Subject Descriptors

D.2.9 [**Software Engineering**]: Management—*Programming teams*

## 1. INTRODUCTION

Biases have been seen to deter meritocracy in offline work groups [6][42]. For years, visible characteristics have been used to differentiate people in all spheres of life ranging from sports [31] to health care [23] to job applications [5]. These biases have started making their presence felt in online environments too [13][14].

OSS started as a merit-based model [33]. It gave rise to terms like ‘code is king’ [4][33]. It was found that social factors influence work-related decisions [37], in addition to the technical factors. In recent years, social work environments like GitHub [3], Bitbucket [1], etc. have concentrated large crowds of developers. These platforms provided transparency and access to developers’ profiles including their contributions and social skills. The increasing level of awareness of demographic attributes of fellow contributors makes it important to understand the reaction of community to this diversity.

Earlier studies on perceived differences in values and norms pointed towards the likelihood of engaging in stereotyping, cliquishness, and conflicts [22]. Recent studies on GitHub have looked into the influence of visible demographic attributes like gender, tenure, etc. on the presence of bias [36] and productivity of teams [39]. Our goal with this work is to understand whether geographical location of developers influences the way their contributions are evaluated. We choose to study the influence of geographical location on the evaluation of contributions for the following reasons: 1) observed impact on work-related decisions in offline groups [6][13], and 2) a reasonable degree of visibility of geographical location in social work environments [38]. Through this study, our aim is to investigate the presence of bias and bridge gaps in perceptions, if any.

To examine bias in online, distributed software development, we leverage GitHub - the largest, most popular online collaborative coding platform. We study pull based development model - one of the most popular model for collaborative development (45% of repositories), which has all characteristics of an online and distributed development. Here, we examine pull request acceptance rate across geographical locations as a proxy to measure bias. The fundamental research question we try to answer is:

> Does the geographical location of the submitter influence pull request acceptance decisions?

One of the key challenges in conducting this study is to identify the presence of bias when developers themselves might not be aware of it. Even when developers are aware of their biases, it is hard to make developers accept it. So, we use a mixed-methods approach. We combine observations from 70,000+ pull requests and 2,500+ survey responses - one of the largest response on open source projects[19], to analyse the influence of geographical location on pull request acceptance decisions.

We quantitatively analyse GitHub projects' data to measure the influence of 1) the geographical location of submitters and 2) the same the geographical location of submitters and integrators on pull request acceptance decisions. We support these observations with the quantitative and qualitative analysis of the survey responses of submitters and integrators on the perceptions of bias. The two roles of developers, submitters and integrators, together present the main stackholders. We combine the results from the three approaches to answer our research question and explain the observed behaviour.

We find that the geographical location of submitters and same geographical location of submitters and integrators explains statistically significant differences in the pull request acceptance decisions. Overall, submitters perceive that they do not experience bias. However, submitters from some geographical locations, such as Brazil and Italy, perceive to have experienced bias, which is more compared to the experiences in other geographical locations. Integrators perceive that they encourage and find it easy to work with submitters from the same geographical location. However, they disagree that submitters from some geographical locations are better at writing pull requests compared to others. Finally, integrators attribute the observed differences in geographical location to language barrier and communication, and not the geographical location of the submitters.

The rest of the paper is organised as follows. In Section 2, we discuss the background and related work followed by the methodology in Section 3. In Section 4, we present the results. Section 5 discusses the threats to validity and Section 6 discusses the implications and conclusions.

# 2. BACKGROUND AND RELATED WORK

## 2.1 Visible demographic attributes and bias

The relationship between visible demographic attributes (like race, gender, etc.) of people and work-related decisions in offline groups has been a subject of study for years [7][5][23]. The influence of these diversities was felt in online communities too. The reaction of the community to these diversities depends on the extent to which these features are salient [39]. OSS started as a merit-based model, [32], however, with the rise of social work environments, like GitHub, developers are somewhat aware of the demographic features (age, gender, ethnicity, etc.) of the fellow developers [38]. This awareness has been been used to form impressions around the history of activities [25]. This is also analysed to understand the influence of gender and tenure diversity on team productivity [39] and the presence of gender bias [36] in GitHub .

## 2.2 Pull based development

GitHub supports two models of collaboration: the shared repository model and the pull based development model. The pull based development model separates development efforts from the decisions to include the submitted code [18]. This separation allowed projects to be more democratic and transparent, which increased participation [26]. Currently, less than half of the collaboratively developed projects exclusively or complementarily use this model [18].

## 2.3 Factors influencing pull request acceptance decisions

Factors influencing pull request acceptance decisions can be broadly classified as developer characteristics, project characteristics, and pull request characteristics. For a developer, reputation (technical and social) is seen to positively influence pull request acceptance decisions [8][20][27][37]. Following technical and social norms are seen to increase the chances of contribution acceptance [20][25][37]. For a project, maturity and popularity is related with lesser chances of pull request acceptance decisions [20][37]. Also, the nature of the pull request, measured as its size (source code churn), quality (including test cases) and uncertainty associated with it (amount of discussion) influence the chances of contribution acceptance [20][37][41].

We control for the confounding effects of factors identified in existing studies and use this information to understand the influence of geographical location on contribution acceptance decisions and the perceptions built around it.

# 3. METHOD

We use a mixed-methods approach [12][15] and triangulate our observations by combining GitHub projects data with survey response data. First, we carefully select a dataset of GitHub developers and projects and model the influence of geographical location on the pull request acceptance decision by controlling for confounding effects. Further, we conduct two large-scale surveys of requesters and integrators in GitHub and quantify their perspectives and experiences. Conducting surveys of submitters and integrators help us put in context and understand the observed behaviour. Below, we present a detailed description of our selection procedure, the justifications, data collection procedures, calculations, and analysis methods. A diagramatic presentation of the research method is shown in Figure 1.

## 3.1 GitHub data

### 3.1.1 Feature Selection

Factors influencing pull request acceptance decisions are borrowed from the literature in software engineering [20][37][39] and social sciences [5][7][23][34]. Table 1 presents a comprehensive list of factors that are seen to influence pull request acceptance decisions. In addition to these factors, for our study, we borrow the concept of bias based on geographical location from the social sciences literature [9] and measure it in terms of software engineering data and its associated meta-data. The list of features used in this study is presented in the last column of Table 1.

### 3.1.2 Data collection

We downloaded the GitHub projects' data designed to analyse pull request development model and is made publicly available by Gousios et al. [19]. We enrich the dataset with additional information required for this study from the GHTorrent dataset made available on August 18, 2015 [2]. We also use the GHTorrent dataset for surveys, as we discuss later.

The enriched dataset is a collection of 1,069 projects and 370,411 pull requests developed in Python (357), Java (315),

**Table 1: A comprehensive list of factors influencing pull request acceptance decisions**

| Characteristics | | Measure | | | |
|---|---|---|---|---|---|
| | | Tsay et. al [37] | Gousios et. al [20] | Current study | Representation |
| **Project characteristics** | | | | | |
| Maturity | | Project tenure | - | Project tenure | repo_pr_tenure_mnth |
| Team size | | Count of collaborators | Active core team | Active core team | - |
| Popularity | | Watchers' count | Watchers' count | Watchers' count | repo_pr_popularity |
| Size of code | | sloc | sloc | sloc | sloc |
| Openness | | - | Percentage of external contribution | Percentage of external contribution | perc_external_contribs |
| Test based code quality | | - | Test lines per kloc | Test lines per lloc | test_lines_per_lloc |
| | | - | Test cases per kloc | Test cases per lloc | - |
| | | - | Asserts per kloc | Asserts per lloc | - |
| **Developer's *acquired characteristics*** | | | | | |
| Social skills | Status in community | Followers | Followers | Followers | prs_popularity |
| | Status in project | Project membership | Project membership | Project membership | prs_main_team_member |
| Social norms | Follow the integrator a priori | Follow the integrator prior to contribution | - | Follow the integrator prior to contribution | prs_followed_pri |
| | Follow the repository a priori | Follow the repository prior to contribution | - | Follow the repository prior to contribution | prs_watched_repo |
| Technical skills | Experience/Expertise | - | Previous pull requests | Previous pull requests | prev_pullreqs |
| | | - | Requester success rate | Requester success rate | prs_succ_rate |
| | | Requester tenure | - | Requester tenure | prs_tenure |
| Technical norms | Size of change | src churn | src churn | src churn | src_churn |
| | | Files changed | Files changed | Files changed | files_changed |
| | Test cases | Test inclusion | Test churn | Test inclusion | test_inclusion |
| **Developer's *innate characteristics*** | | | | | |
| Geographical location | | - | - | Country of residence | prs_location |
| Explicit or implicit bias on geographical location | | - | - | Likelihood of PR acceptance on geo loc | measured using prs_location |
| In-group bias | | - | - | Likelihood of PR acceptance on same geo loc | measured using prs_pri_same_location |
| **Pull request characteristics** | | | | | |
| Uncertainty associated with pull requests | | Comments count | Comments count | Comments count | num_issue_comments |

Ruby (359), and Scala (38). These carefully selected projects represent top 1% of the projects developed by using pull request development model. It combines GHTorrent data with project repositories data and provides a list of features seen to influence pull request development. For this study, we augment this dataset with pull request life cycle information, participants' demographic information, and measures of social norms that influence pull request acceptance decision. A description of the procedure to extract the above-mentioned factors from the GHTorrent dataset follows.

### *3.1.3 Features*

**Participants' demographics**

We measure geographical location of developers as the location (or country of residence) specified by developers in their GitHub profiles. Following two factors motivates our choice of country of residence as a measure of geographical location. First, the perceptions of fellow developers around geographical location are framed in terms of the location specified in GitHub profile. Second, we believe that work habits and cultural environment, specific to the location of current residence of a developer, may explain differences in the behaviour. One may argue that these differences are caused by a combination of current and attenuating effects of past locations of residence. However, for simplicity, we study the differences in pull request acceptance decisions in terms of current location only.

In GitHub, mentioning the location is optional and is specified in a free-form text. Thus, developers can write 'US', 'United States', 'XYZ Apartments, New York', etc. to refer to the same location. To identify the geographical location of developers irrespective of the format of the location, we use 'countryNameManager' script used by Vasilescu et. al [39]. This script uses free-form text to identify the geographical location of GitHub users, who choose to mention it. Further, to identify the geographical location of developers who did not mention it explicitly, we proposed some heuristics.

The proposed heuristic uses the data points for which geographical location is identified using 'countryNameManger' as training data. It then uses domain names of email addresses and affiliation of developers to identify their geographical location. The underlying principle used by the heuristics is that affiliation and domain name can be used to localise country of residence. For instance, affiliation to 'Peking University' maps the country of residence to China. However, this approach is prone to false-positives. To minimise false positives, we choose one-to-one mapping between predictors and geographical location (exclude one-to-many) and set the threshold for inclusion as 20. Thus, if in our training dataset company name 'Peking University' maps to China only and has at least 20 data points to support it, we map 'Peking University' to China. This information is used to identify the geographical location of developers for whom we cannot identify the geographical location using 'countryNameManager' script and who have mentioned company name or email address. This way, we identified the geographical location of 149,268 developers (out of 541,685) who participated in pull request based development. Here, participants refer to both submitters and integrators.

**Life cycle of a pull request**

A pull request is opened by a requester and is in state 'open'. Integrators (core team developers) review the pull request. The integrators evaluate the opened pull request

**Table 2: Pull request data generated**

| PR ID | State: Open | State: Merge | State: Close | Status | Same Location |
|---|---|---|---|---|---|
| 1 | Dev1 (India) | - | - | Open | - |
| 2 | Dev1 (India) | Dev2 (India) | - | Merged | Yes |
| 3 | Dev1 (India) | Dev3 (US) | Dev3 (US) | Merged | No |
| 4 | Dev1 (India) | - | Dev4 (France) | Not Merged | No |
| 5 | Dev1 (India) | Dev1 (India) | Dev1 (India) | Merged | Self |

and decide to 1) merge the suggested change in the original code with a state 'merge', followed by state 'close' or 2) close it directly without merging it with a state 'close'. A pull request can be in state 'open', 'merged', or 'not merged'. A pull request may get re-opened multiple times. However, here we focus on the pull request lifecycle starting from the time when the pull request was opened first till the time it gets closed the first time.

We select pull requests which are 'merged' and 'not merged', and exclude 'open' pull requests. We construct pull request life cycle and append it with the developer of the action, geographical location of developer and others. Here, developers who opened pull requests are marked as 'submitter'. Developers who merged the pull requests are termed as 'merger' and developers who closed the pull requests as 'closer'. Together, 'merger' and 'closer' are integrators and are core team of the project. The final data from this exercise looks like the one in Table 2.

In Table 2, submitter Dev1 from India interacts with integrators Dev1, Dev2, Dev3, and Dev4 from various geographical locations to get her pull requests reviewed. The outcome of the pull request review (status) and the relationship between the geographical location of submitters and integrators (same geographical location) is identified.

### *3.1.4 Pull request project sample*

The geographical location of submitters follows a highly skewed distribution (kurtosis: $\gamma$=98.2). So, to ensure that we have diverse geographical locations and significant pull request counts for each location, we select geographical locations, which represent atleast 1% of the total pull requests of the GitHub data. Thus, the United States (38%), United Kingdom (8%), Germany (6%), France (5%), Canada (4%), Japan (3%), Brazil (3%), Australia (2%), Russia (2%), Netherlands (2%), China (2%), Spain (2%), India (2%), Switzerland (1%), Sweden (1%), Italy (1%), and Belgium (1%) with at least 1% of total pull requests are selected for analysis. These selected 17 geographical location represent approximately 83% of the total developer population for whom we were able to identify geographical location.

We observed that the submitters themselves merge a significant fraction of pull requests on GitHub. An analysis of 113,191 pull requests in the enriched dataset showed that other developers integrate 63% of the pull requests and the submitters themselves integrate the remaining 37%. Similar statistics are observed in the GHTorrent dataset (44% merged by self, 56% by others). So, by selecting pull requests from submitters from the selected 17 geographical locations, eliminating pull requests merged by self and incomplete data, we are left with 70,740 pull requests for analysis.

### *3.1.5 Statistical methods*

In this study, a pull request is the base unit of analysis. To understand the influence of geographical location of a submitter and its interaction with the geographical location of an integrator on the pull request acceptance decision, we test two hypotheses.

$H1_0$ There are no differences in the pull request acceptance decisions based on the geographical location of submitters.

$H1_a$ There are differences in the pull request acceptance decisions based on the geographical location of submitters.

$H2_0$ There are no differences in the pull request acceptance decisions based on the same geographical location of submitters and integrators.

$H2_a$ There are differences in the pull request acceptance decisions based on the same geographical location of submitters and integrators.

To test the two hypotheses, we use regression modeling. We model the pull request acceptance decisions as a binary classification problem. Specifically, we use logistic regression, as implemented in R [21][29]. We measure statistical significance at a p-value $<=$0.05, size of change as log odds, and the impact as the percentage of deviance as used in other studies [30]. Though this measure provides an interpretation similar to the percentage of the total variance explained by least square regressions, the two measures are not the same [10].

At a finer level of detail, we compute a pairwise correlation of continuous variables and note down the two highly correlated variables. We consider two variables as highly correlated when their correlation coefficient is greater than 0.7 [16]. Similarly, to measure the relationship between two categorical variables, we use a chi-square test for dependence for significance and measure its effect size using Cramer's V [11]. For categorical variables, we consider that the relationship between two variables is strong if it exceeds 0.7 [24]. We identify and note strongly correlated categorical variables too. We model the relationship of a set of predictors against response variable. Next, to stabilise the variance, we log-transform the independent count variables. We verify this by using AIC and Vuong test for non-nested models [40] to compare the transformed and non-transformed data. To check that multicollinearity is not an issue, we compute the Variance Inflation Factor (VIF). Any VIF value greater than 5 is considered to indicate multicollinearity, as used in various studies [10]. At this step, we eliminate highly correlated variables that cause multicollinearity. We measure the fitness of model using Area under Curve (AUC). The value of AUC should be greater than 0.5 for the model to be acceptable[1].

Once the model is built, we read the coefficients of logistic regression as the expected changes in the log of responses

[1] https://www.kaggle.com/wiki/AUC

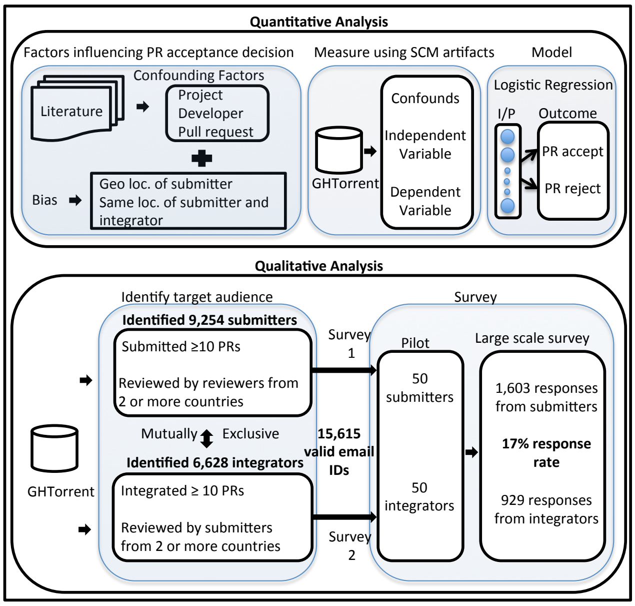


**Figure 1: Research Method: Mixed-Methods Approach**

for a unit change in the predictor variable while keeping all other predictor variables at a constant value. So, for continuous predictor variables, one unit change in its value is associated with an exponent of the coefficient change in the response variable. The interpretation is slightly different for categorical variables. To analyse the effect size of statistically significant features, we use dummy or treatment code to compare a base level with treatments. Specifically, to measure the influence of geographical location, we choose the United States, with the majority of developers, as the base level. Similarly, to test for bias based on the interactions between the geographical location of submitters and integrators, we choose 'different geographical location' as the base level.

## 3.2 Survey

Unlike the quantitative analysis of GitHub projects data, which we conducted on a small, carefully selected set of pull requests to provide a controlled environment, we conduct a large-scale survey of a wide range of developers. We designed two surveys - one from the perspective of submitters and the other from the perspective of integrators. The choice of conducting two separate surveys helps us understand the complete picture of perspectives and experiences.

### 3.2.1 Design

Since to goal of conducting two surveys is to learn from the developer community rather than a few individuals, we designed online surveys [17] [35]. We did not give any monetary incentives. The survey was designed to take a maximum of 7 minutes of survey respondents and was active for three weeks.

There were two rounds for each survey. First, we identify 50 developers each in the role of the submitter and the integrator and sent them the survey as part of the pilot study. Based on their feedback and refinement, we sent the main survey to all submitters and integrators identified to answer our surveys. We explicitly informed our prospective survey respondents that the results of the study would be anonymous to ensure that developers share their true opinion and not try to present themselves in good light. The details of each survey are given below:

**Integrators** We asked our survey respondents four types of questions. First, we asked them a few demographic questions to get an understanding of the diversity and representativeness of the survey responses. This was followed by questions to understand their perspectives on the presence of bias. These questions are

1. Level of awareness of the demographic features of the submitters they work with.

2. Perceived relevance of the importance of developers' characteristics.

3. Explicit perceptions of bias based on the geographical location of submitters and its interaction with the geographical location of integrators.

In total, we asked 12 questions, out of which there were 8 multiple-choice questions, 3 Likert scale questions and 1 open-ended question. The aim of the open-ended question was to discover factors other than the ones mentioned in the survey.

**Submitters** We asked our survey respondents two types of questions. Similar to the survey for integrators, we started with asking a few demographic questions. This was followed by questions to understand the following:

1. Their understanding of the importance of factors influencing pull request acceptance decision of integrators.

2. Their personal experiences with bias.

In total, we asked them 11 questions, out of which there were 8 multiple choice questions, 2 Likert scale questions and 1 open ended question.

Throughout the two surveys, we used 5-point Likert scale for the study with one exception. To understand the perceptions of submitters on "The important of factors influencing pull request acceptance decision of integrators", we provided an additional choice - 'I don't know'. This choice was provided to account for cases when submitter doesn't know the perceived importance of the factor.

### 3.2.2 Identify Survey Respondents

From the GHTorrent dataset, we extracted developers' geographical location information, role, and contribution count to identify a list of survey respondents for this study. We selected all submitters who have submitted at least 10 pull requests that were reviewed by integrators from 2 or more geographical locations. Here, the count of pull requests submitted does not include the pull requests closed by self. Similarly, we identified integrators who reviewed at least 10 requests from submitters from 2 or more geographical location. Here also the count of pull requests excluded the pull requests merged by self. These selection criteria ensure the following:

1. *Candidacy of the developer to answer our survey questions*: The choice of two or more geographical locations ensures that the developer has experience working with

developers from diverse geographical locations. Such developers can help us understand the influence of geographical location on the pull request acceptance decisions.

2. *Reasonable experience and diversity of respondents*: The choice of working on at least 10 pull requests not closed by self, ensure that the respondent has reasonable experience working with the pull-based development model. It ensures that we include a wide range of developers for the study.

Developers can play the roles of submitter and integrator simultaneously. To select unique respondents per survey, we select developers for the role they worked the most. So, a developer who wrote 100 pull requests and reviewed 150 pull requests is considered for the role of the integrator. Following this approach, we identified 6,628 integrators and 9,254 submitters. Out of these 15,882 prospective respondents, only 15,615 respondents had a valid email address.

We sent customised emails to all identified prospective survey respondents. We informed the developers about their contribution in terms of the approximate count of pull requests they worked on and an approximate count of the geographical locations with which they collaborated. This customised report generated interest in the survey, which was reflected in the 500+ email responses received and 2,532 survey responses - one of the largest survey responses with 17% response rate (excluding 818 messages which failed to deliver).

#### 3.2.3 Data analysis

We identified all complete survey responses, preprocessed it and got it into a form usable for analysis. To hypothesis test the research questions, we converted the ordinal data to its nominal equivalent and conducted chi-square test.

We started with some basic understanding of the diversity of responses in terms of respondents' demographics. This is followed by the perceived importance of the geographical location of requester on the pull request acceptance decision across the two roles. We additionally analysed the level of awareness of integrators regarding submitters' demographics. Finally, to understand the perceptions on bias based on geographical location, we asked submitters and integrators different questions.

- *Question for submitters* Did you experience bias based on your geographical location in getting your pull requests accepted?

- *Questions for integrators* How much do you agree with the following statement?

  - It is easy to work with developers from the same geographical location.
  - I encourage developers from my geographical location to contribute.
  - Developers from some geographical locations are better at writing pull requests relative to others.

To understand the perceptions of developers, we hypothesis tested the response as follows

$H_0$ The experiences on bias are of equal proportions.
$H_a$ There are unequal proportions of experiences on the perceptions on bias.

If at a 0.05 significance level null hypothesis is rejected, we measure the effect size as percentage difference. Next, we see if there are significant difference in perceptions on bias across geographical locations. If the differences are significant, we measure differences in experiences around bias using logistic regression. For this we select geographical locations for which we received at least 10 responses.

We codified open-ended survey response to get an understanding of the justification of integrators around the presence of bias based on geographical location. The first author started with identifying themes by analysing first 100 comments. The themes were the top three suggestions that integrator would like to give to submitters to improve the chances of their pull requests getting accepted. These themes, along with any other theme that evolved in time were codified for all open-ended responses. One other author then verified this. These suggestions conveyed the expectations of integrators from submitters and also pointed towards the possible explanation of observed differences in perceptions.

## 4. RESULTS

### 4.1 Data Analysis

To answer our research question, we test Hypothesis 1 and Hypothesis 2 using GitHub projects' data. We model the influence of geographical location on pull request acceptance decisions using logistic regression while controlling for the effect of confounding factors. The details of the model are shown in Table 3.

*H1: There are differences in the pull request acceptance decisions based on the geographical location of submitters.*

In Model 1 of Table 3, we see that increase in maturity, popularity, size of the code, and openness to external contribution reduces the chances of pull request acceptance decision while a good tested code - an indicator of code quality increases it. We see that increase in technical skills of the submitters and abiding by technical norms are seen to increase the chances of pull request acceptance while increase in experience decreases it. We also notice that submitters with a high social reputation and those who follow social norms are more likely to get their pull requests accepted. Finally, the uncertainty associated with the pull request influences the chances of pull requests getting accepted. The influence of the above-mentioned factors are already known to the software engineering community [19][37]. Controlling for the effect of these factors in Model 1 in Table 3, we see that the geographical location of submitters explains significant differences in the chances of the pull request acceptance. The deviance explained by each factor in shown in Table 4.

Our observations support our hypothesis that there are differences in the pull request acceptance decision based on the geographical location of submitters. Next, we identify the differences in the pull request acceptance decision location-wise. We see the sign and magnitude of the estimate for a given geographical location relative to the base

| | Model 1 | Model 2 |
|---|---|---|
| (Intercept) | 2.82 (0.14)*** | 2.61 (0.14)*** |
| Control variables | | |
| repo_pr_tenure_mnth | −0.01 (0.00)*** | −0.01 (0.00)*** |
| repo_pr_popularity | −0.00 (0.00)*** | −0.00 (0.00)*** |
| perc_external_contribs | −0.01 (0.00)*** | −0.01 (0.00)*** |
| test_lines_per_lloc | 0.00 (0.00)*** | 0.00 (0.00)*** |
| log(sloc + 1) | −0.06 (0.01)*** | −0.06 (0.01)*** |
| prs_tenure | −0.00 (0.00)*** | −0.00 (0.00)*** |
| log(prev_pullreqs + 1) | 0.17 (0.01)*** | 0.17 (0.01)*** |
| prs_succ_rate | 0.01 (0.00)*** | 0.01 (0.00)*** |
| test_inclusion1 | 0.26 (0.03)*** | 0.26 (0.03)*** |
| log(src_churn + 1) | −0.06 (0.01)*** | −0.06 (0.01)*** |
| log(files_changed + 1) | 0.01 (0.02) | 0.01 (0.02) |
| prs_main_team_member1 | 0.06 (0.07) | 0.05 (0.07) |
| log(prs_popularity + 1) | 0.06 (0.01)*** | 0.07 (0.01)*** |
| log(num_issue_comments + 1) | −0.25 (0.01)*** | −0.24 (0.01)*** |
| prs_watched_repo1 | 0.04 (0.03) | 0.05 (0.03) |
| prs_followed_pri1 | 0.11 (0.03)** | 0.10 (0.03)** |
| prs_locationunited kingdom | 0.13 (0.04)** | 0.20 (0.04)*** |
| prs_locationgermany | −0.25 (0.04)*** | −0.16 (0.05)*** |
| prs_locationfrance | 0.02 (0.06) | 0.11 (0.06) |
| prs_locationcanada | 0.12 (0.07) | 0.22 (0.07)** |
| prs_locationjapan | 0.25 (0.08)*** | 0.34 (0.08)*** |
| prs_locationbrazil | −0.27 (0.06)*** | −0.19 (0.07)** |
| prs_locationaustralia | 0.05 (0.07) | 0.14 (0.07) |
| prs_locationnetherlands | 0.26 (0.09)** | 0.36 (0.09)*** |
| prs_locationchina | −0.39 (0.09)*** | −0.27 (0.10)** |
| prs_locationrussia | −0.06 (0.07) | 0.04 (0.07) |
| prs_locationspain | 0.08 (0.10) | 0.15 (0.10) |
| prs_locationindia | 0.02 (0.07) | 0.12 (0.07) |
| prs_locationswitzerland | 0.38 (0.11)*** | 0.46 (0.11)*** |
| prs_locationsweden | −0.21 (0.09)* | −0.10 (0.09) |
| prs_locationbelgium | 0.09 (0.12) | 0.18 (0.12) |
| prs_locationitaly | −0.31 (0.08)*** | −0.21 (0.09)* |
| prs_pri_same_location1 | | 0.18 (0.03)*** |
| AIC | 49231.59 | 49198.20 |
| BIC | 49534.09 | 49509.87 |
| Log Likelihood | −24582.80 | −24565.10 |
| Deviance | 49165.59 | 49130.20 |
| Num. obs. | 70740 | 70740 |

***$p < 0.001$, **$p < 0.01$, *$p < 0.05$

**Table 3: Logistic regression model of factors influencing pull request acceptance decision [AUC: 0.7]**

| | Df | Deviance | Resid. Df | Resid. Dev | Pr(>Chi) |
|---|---|---|---|---|---|
| NULL | | | 70739 | 52354.02 | |
| repo_pr_tenure_mnth | 1 | 509.46 | 70738 | 51844.56 | 0.0000 |
| repo_pr_popularity | 1 | 175.49 | 70737 | 51669.07 | 0.0000 |
| perc_external_contribs | 1 | 321.47 | 70736 | 51347.59 | 0.0000 |
| test_lines_per_lloc | 1 | 47.84 | 70735 | 51299.75 | 0.0000 |
| log(sloc + 1) | 1 | 30.85 | 70734 | 51268.90 | 0.0000 |
| prs_tenure | 1 | 25.46 | 70733 | 51243.45 | 0.0000 |
| log(prev_pullreqs + 1) | 1 | 1014.64 | 70732 | 50228.81 | 0.0000 |
| prs_succ_rate | 1 | 389.62 | 70731 | 49839.19 | 0.0000 |
| test_inclusion | 1 | 22.73 | 70730 | 49816.46 | 0.0000 |
| log(src_churn + 1) | 1 | 164.02 | 70729 | 49652.44 | 0.0000 |
| log(files_changed + 1) | 1 | 0.11 | 70728 | 49652.34 | 0.7444 |
| prs_main_team_member | 1 | 2.62 | 70727 | 49649.72 | 0.1055 |
| log(prs_popularity + 1) | 1 | 57.46 | 70726 | 49592.26 | 0.0000 |
| log(num_issue_comments + 1) | 1 | 273.45 | 70725 | 49318.81 | 0.0000 |
| prs_watched_repo | 1 | 2.82 | 70724 | 49315.99 | 0.0931 |
| prs_followed_pri | 1 | 6.79 | 70723 | 49309.20 | 0.0092 |
| prs_location | 16 | 143.61 | 70707 | 49165.59 | 0.0000 |
| prs_pri_same_location | 1 | 35.39 | 70706 | 49130.20 | 0.0000 |

**Table 4: Deviance explained by factors influencing pull request acceptance decision**

level - the United States. The coefficients for submitters' geographical location can be divided into three categories: statistically insignificant coefficients, positive coefficients, and negative coefficients. For 7 out of the 17 geographical locations under analysis, the results are statistically insignificant, that is, there are no differences in the chances of contribution acceptance decision compared to the United States. For geographical locations, where coefficients are positive, the chances that their pull requests get accepted are higher than that of the United States. Similarly, geographical locations for which coefficients are negative have lower chances of getting their pull requests accepted. Thus, France, Australia, Russia, Spain, India, Sweden and Belgium observe no differences in getting their pull requests accepted compared to the United States. United Kingdom (22%), Canada (25%), Japan (40%), Netherlands (43%), and Switzerland (58%) have higher chances of getting their pull requests accepted. Germany (15%), Brazil (17%), China (24%), and Italy (19%) have lower chances of getting their pull requests accepted. In Table 4, we see that the geographical location of submitters explains a small, yet significant percentage of the total deviance.

*H2: There are differences in the pull request acceptance decisions based on the same geographical location of submitters and integrators.*

To test this hypothesis, we control for the effect of abovementioned factors including the geographical location of submitters. In Model 2 of Table 3, we see that the same geographical location of submitters and integrators, compared to different geographical locations of submitters and integrators, has statistically significant influence on pull request acceptance decisions. This observation supports our hypothesis 2. We see that controlling for the effects of other factors, when submitters and integrators are from the same geographical location there is 19% more chances that the pull requests will get accepted compared to when submitters and integrators are from different geographical location. Further, in Table 4, we see that the same geographical location of submitters and integrators explains a small, yet significant percentage of the total deviance.

## 4.2 Survey

### 4.2.1 Submitters

We received 1,603 complete responses from submitters. These responses present perspectives of a wide range of submitters from 76 countries with different age [21-30 years (47%), 31-40 years (42%), 41-50 years (7%) and others] , gender [male (98%), female (2%)], experience in OSS [1-2 years (10%), 3-6 years (53%), 7-10 years (19%), more than 10 years (17%) and others], job [Industry (62%), Academia (5%), Freelance (7%) and others], and role [source code contributor (49%), owner (47%) and others]. We analyse the perceptions of these submitters to answer our hypothesis H3.

*H3: Submitters perceive bias based on their geographical location.*

There are statistically significant differences in the perceptions of submitters on the presence of bias. Using chi-square test of independent at a significance level of 0.05 (X-squared = 1448.743, df = 1, p-value < 2.2e-16) we found that 97% more developers feel that did not experience bias compared

| | Bias |
|---|---|
| (Intercept) | 4.84 (0.50)*** |
| prs_locationGermany | 0.01 (1.12) |
| prs_locationUnited Kingdom | 16.73 (2908.74) |
| prs_locationFrance | 16.73 (3116.19) |
| prs_locationCanada | −0.74 (1.13) |
| prs_locationRussia | 16.73 (3906.35) |
| prs_locationBrazil | −2.60 (0.69)*** |
| prs_locationAustralia | 16.73 (4015.38) |
| prs_locationIndia | −1.20 (1.13) |
| prs_locationNetherlands | 16.73 (4941.18) |
| prs_locationItaly | −2.13 (0.89)* |
| prs_locationSweden | 16.73 (5250.30) |
| prs_locationSpain | 16.73 (5524.41) |
| prs_locationJapan | −2.80 (0.79)*** |
| prs_locationNorway | −1.75 (1.14) |
| prs_locationPoland | −1.75 (1.14) |
| prs_locationSwitzerland | 16.73 (6379.04) |
| prs_locationUkraine | −2.07 (1.15) |
| prs_locationDenmark | 16.73 (7812.70) |
| prs_locationCzech Republic | 16.73 (8107.62) |
| prs_locationFinland | 16.73 (8107.62) |
| prs_locationPortugal | −2.35 (1.16)* |
| prs_locationBelgium | 16.73 (8438.68) |
| prs_locationArgentina | −2.54 (1.16)* |
| prs_locationAustria | 16.73 (8813.91) |
| prs_locationIreland | 16.73 (8813.91) |
| prs_locationSingapore | 16.73 (9244.11) |
| AIC | 236.25 |
| BIC | 378.45 |
| Log Likelihood | −91.12 |
| Deviance | 182.25 |
| Num. obs. | 1432 |

$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$

**Table 5: Submitters' perception on bias - location-wise [AUC: 0.9]**

to those who feel that they experienced it. Our hypothesis H3, that is, submitters perceive bias based on their geographical location is rejected. Further analysis showed that the differences in the perceptions on bias across geographical location are statistically significant (X-squared = 72.1569, df = 26, p-value = 3.203e-06), that is, there are differences in the perceptions on bias based on the geographical location of submitters. To tease out the individual effects, we modeled the perceptions on the experience of bias location-wise using logistic regression. An analysis of 27 geographical locations from which we received at least 10 survey responses show that more submitters from some geographical locations perceive the presence of bias compared to others. We found that the perceptions on the presence of bias are more for submitters from Brazil (93%), Italy (87%), Japan (94%), Portugal (90%), and Argentina (93%) compared to other geographical locations (refer Table 5).

### 4.2.2 *Integrators*

We received 929 complete responses from integrators from 61 different geographical locations. These respondents were diverse in terms of age [21-30 years (38%), 31-40 years (46%), 41-50 years (12%)], gender [male (97%), female (3%)], experience in OSS [1-2 years (6%), 3-6 years (43%), 7-10 years (24%), more than 10 years (26%) and others] , job [Industry (66%), Freelance (10%), Academia (6%) and others] and role [Owner (88%) and Source code (11%)]. We started the survey by understanding their perceptions on the level of awareness of the geographical locations of the submitters they work with. 43% of integrators say that they are rarely aware of the geographical location of submitters they work with. This is followed by 28% of integrators who feel that they sometimes know geographical location, followed by 16% who never know, 10% often and only 3% always. This was followed by a direct question to understand their perceptions on the importance of geographical location on pull request acceptance decision for which we tested hypothesis H4.

*H4: Integrators perceive that the geographical location of submitters is important on their pull request acceptance decisions.*

88% more developers perceive that the geographical location of submitters is unimportant than those who considered it important (X-squared = 705.8317, df = 1, p-value < 2.2e-16) and this was consistent across all geographical locations (X-squared = 12.958, df = 16, p-value = 0.6758). Thus, our observation refutes our hypothesis H4. This was followed by more specific questions to understand the influence of geographical location.

*H5: Developers from some geographical locations are better at writing pull requests relative to others.*

43% of the respondents choose to not comment on this question. An analysis of the integrators who expressed their opinion show with statistical significance that 52% more integrators disagree than those who agree that developers from some geographical locations are better at writing pull requests than others (X-squared = 138.1231, df = 1, p-value < 2.2e-16). This refutes our hypothesis H5 that developers from some geographical locations are better at writing pull requests relative to others. On digging deeper, we found that the perceptions are similar across geographical locations (X-squared = 18.3874, df = 11, p-value = 0.07302), with an exception of India, where integrators felt that developers from some geographical locations are better at writing pull requests. For the rest of the geographical locations, for every 1 developer that feels that there are differences in the abilities of submitters to write pull requests based on geographical locations, there are 4 developers that disagree to it. On a contrary, half of the integrators from India agree and the other half disagrees on it (refer Table 6).

*H6: Integrators perceive that it easy to work with submitters from the same geographical location.*

55% of the integrators' population preferred not to answer this question. The subset of integrators who choose to express their opinion felt that it easy to work with submitters from their own geographical location with statistical significant results (X-squared = 80.597, df = 1, p-value < 2.2e-16). There were 45% more integrators who felt that it is easy to work with submitters from their own geographical location. This supports our hypothesis H6. Further, we examined perceptions across geographical locations and found a different perspective (X-squared = 24.4768, df = 8, p-value = 0.001906). On one side, countries like the United States (8 out of 10) and United Kingdom (6 out of 10) agree that it is easy to work with developers from the same geographical location. On the other side, integrators from Germany and India disagree 6 out of 10 and 8 out of 10 times respectively. In short, a majority of the geographical locations analysed feel that it is easy to work with developers from the same geographical location with few exceptions (refer Table 7).

*H7: Integrators encourage submitters from their geographical location to participate.*

| | Better loc on PR |
|---|---|
| (Intercept) | 1.30 (0.16)*** |
| pri_locationUnited Kingdom | 0.84 (0.55) |
| pri_locationGermany | 0.22 (0.45) |
| pri_locationFrance | −0.52 (0.52) |
| pri_locationCanada | −0.04 (0.59) |
| pri_locationSweden | 1.19 (1.05) |
| pri_locationAustralia | −0.71 (0.58) |
| pri_locationNetherlands | −0.09 (0.68) |
| pri_locationIndia | −1.45 (0.58)* |
| pri_locationBrazil | 0.90 (1.07) |
| pri_locationSwitzerland | 1.27 (1.05) |
| pri_locationNorway | −0.45 (0.71) |
| AIC | 440.03 |
| BIC | 488.51 |
| Log Likelihood | −208.01 |
| Deviance | 416.03 |
| Num. obs. | 420 |

$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$

**Table 6: Integrators' perception on pull request quality and geographical location [AUC: 0.6]**

| | In group ease |
|---|---|
| (Intercept) | 1.50 (0.19)*** |
| pri_locationUnited Kingdom | −1.10 (0.45)* |
| pri_locationGermany | −1.76 (0.46)*** |
| pri_locationFrance | −0.55 (0.56) |
| pri_locationCanada | 1.14 (1.05) |
| pri_locationSweden | 15.07 (665.51) |
| pri_locationAustralia | 0.11 (0.80) |
| pri_locationIndia | −2.89 (0.81)*** |
| pri_locationBrazil | −1.50 (0.84) |
| AIC | 323.24 |
| BIC | 356.99 |
| Log Likelihood | −152.62 |
| Deviance | 305.24 |
| Num. obs. | 314 |

$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$

**Table 7: Integrators' perception on ease to work with the same location developers [AUC: 0.7]**

Like previous questions, 53% of the integrators preferred not to answer this question. From the remaining, 26% more integrators agree that they encourage developers from their geographical location to participate with statistically significant results (X-squared = 27.8244, df = 1, p-value = 1.328e-07). This supports our hypothesis H7. Further, we see that perceptions on this depend on the geographical location of integrator (X-squared = 24.3008, df = 8, p-value = 0.00204). Unlike other geographical locations, integrators in United Kingdom (62%), Germany (69%) and Sweden (77%) are less likely to encourage developers from their geographical location to participate (refer Table 8).

### 4.3 Open-ended survey responses

We received 639 open-ended survey responses from integrators where they talked about factors that influence their acceptance decisions. From manually coding these open-ended responses, six themes came up: *technical skills (47%), relevance of requested feature (12%), communication skills (23%), behaviour (11%), trust (4%), and pro activeness (2%).* We went deeper into non-technical aspects to uncover the expectations of integrators that may possibly explain the perceived differences in pull request acceptance decision. One key observation from these responses was the ability to communicate. Integrators mentioned that they form an impression about the pull request based on the description of the

| | In group encourage |
|---|---|
| (Intercept) | 0.81 (0.16)*** |
| pri_locationUnited Kingdom | −1.00 (0.40)* |
| pri_locationGermany | −1.18 (0.42)** |
| pri_locationFrance | −0.19 (0.50) |
| pri_locationCanada | 0.20 (0.61) |
| pri_locationSweden | −1.50 (0.63)* |
| pri_locationNetherlands | −0.25 (0.65) |
| pri_locationIndia | 0.17 (0.70) |
| pri_locationBrazil | 15.76 (758.80) |
| AIC | 399.35 |
| BIC | 433.04 |
| Log Likelihood | −190.67 |
| Deviance | 381.35 |
| Num. obs. | 312 |

$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$

**Table 8: Integrators' perception on encouraging developers from same geographical location [AUC: 0.6]**

pull request. Integrator [R496] mentioned that *"bad or misleading title influence her chances of accepting the pull request"*.

They also used the ability to communicate to justify their perceptions on the importance of geographical location of submitters in explaining pull request acceptance decision. Integrator [R289] stated that *"Nationality really only plays a role in my ability to understand the written communication. Being distributed makes communication more important, so if I cannot understand the requester, I'm much less likely to accept the request."*. Further, integrator [R201] added that *"There's a possibility of language bias - if the pull request isn't well-written (which is often the case when English is not the PR author's first language) I may be more hesitant, but usually because of a fear of misunderstanding."*. Integrator [R883] even goes on to add that *"I often reject pull requests that add in code that has misspelled words, poor grammar, etc., and ask the contributor to fix those before merging."*. In addition to this, integrators also mentioned the importance of behaviour, trust and pro activeness of submitters on their pull request acceptance decision. Integrator [R738] said that the tone of the pull request's body is important for her. She said that *"I don't want to work together if the person is rude."*. Integrator [R773] even goes on to say that *"If they<submitters> are rude, their pull request is rejected, even if the code quality is great."*. Besides these, integrators place their bet on the submitters they trust. Integrator [R50] stated that *"People that have submitted good PRs in the past I almost blindly merge"*. Finally, integrators appreciate the willingness of the submitters to quickly make desired changes to improve contribution. Integrators ruled out the presence of explicit bias based on the geographical location in favour of the valuable contribution they receive for their projects. In this context, integrator [R661] quoted that *"I see PRs as a favor to me, so I tend to take it seriously to process PRs ASAP and treat requesters well."*.

## 5. THREATS TO VALIDITY

### 5.1 Internal Validity

*Data accuracy* The accuracy of the results of the study depends on the accuracy of the data on which it is built. We have used GHTorrent data, which has been extensively used in several prior studies. So, we believe that it should help that it is a precise dataset.

*Language bias* The surveys floated for this study were written in English. We justify our choice by conducting a pilot study where an equal number of surveys were sent out in English and French to developers in France. We received similar response rates from both. This is intuitive as developers who use GitHub must be aware of basic English used in the survey. Still there is a possibility that the choice of English biased the response rates from some geographical locations.

*Researcher bias* To prevent researcher's bias on the articulation of our questions, we got our survey questions validated by a wide range of people including survey design experts even before making it public. These people checked the language of the questions for ambiguity and presence of bias.

*Research reactivity* The tendency of the respondents to appear in positive light may influence the results.

*Non-response bias* It is possible that the developers who did not responded to the survey may have different insights. However, we do not see this as a big concern as we received survey responses from more than 76 geographical locations.

## 5.2 External Validity

*Generalisability* The quantitative analysis present in this study is built on small, carefully selected projects. These projects may not be representative of all the collaborative developed projects. We try to address this concern by combining the results of the quantitative data with large-scale survey responses. Further, while we conducted this survey on GitHub - biggest code hosting sites featuring pull-based development model, we believe that similar experiments must be conducted on other platforms, like Bitbucket, for generalisability.

# 6. IMPLICATIONS AND CONCLUSIONS

We analysed 70,000+ pull requests from GitHub and 2,500+ survey responses to study the presence of bias based on the geographical location. Data analysis show that the geographical location of submitters significantly influences the pull request acceptance decisions. Compared to the United States, submitters from United Kingdom (22%), Canada (25%), Japan (40%), Netherlands (43%), and Switzerland (58%) have higher chances of getting their pull requests accepted. However, submitters from Germany (15%), Brazil (17%), China (24%), and Italy (19%) have lower chances of getting their pull requests accepted. We observed that the same geographical location of submitters and integrators increases the chances of pull request acceptance by 19% compared to when submitters and integrators are from different geographical locations.

> Data analysis suggests the presence of bias based on the geographical location.

Survey responses from submitters show that overall submitters do not perceive that they experience bias. We found that 97% more submitters feel that they did not experience bias compared to those who felt that they experienced it. However, submitters from some geographical locations perceive to have experienced bias, which is more compared to other geographical locations. We found that the perceptions on the presence of bias are stronger for submitters from Brazil (93%), Italy (87%), Japan (94%), Portugal (90%), and Argentina (93%) when compared to other geographical locations.

> Submitters from some geographical locations perceive to experience bias.

The observations of data analysis and the perceptions of submitters match. Submitters from Brazil and Italy perceive to experience bias more than other geographical locations. The same is observed in our analysis of GitHub data that developers from Brazil and Italy have lower chances of getting their pull requests accepted compared to other geographical locations.

> Perceptions of submitters on the presence of bias based on the geographical location are in agreement with the data analysis.

53% more integrators perceive that they encourage developers from their country to participate. Also, 8 out of every 10 integrators feel that it is easy to work with submitters from the same geographical location. However, they do not feel that developers from some geographical locations are better at writing pull requests compared to others, with an exception of India. For every 1 developer who feels that there are differences in the abilities of submitters to write pull requests based on geographical location, there are 4 developers that disagree to it. However, in India half of the integrators agree and the other half of the integrators disagree to it.

> Integrators perceive that they are not biased in evaluating submitters.

Open-ended survey responses from integrators present another perspective to the observation. In open-ended survey responses integrators suggest that the observed differences may be explained in terms of language barriers and the ability to communicate, and not necessarily bias based on the geographical location of submitters.

> Integrators think that factors relating to the geographical location and not necessarily the geographical location may influence their pull request acceptance decisions.

> There exists a bias blind spot [28] - a cognitive bias, as integrators perceive the absence of bias and submitters perceive to experience it.


# 7. ACKNOWLEDGMENTS

The authors would like to acknowledge all survey-respondents for their valuable time and feedback.

## 8. REFERENCES


[1] 2015. Bitbucket. (2015). https://bitbucket.org.

[2] 2015. GHTorrent. (2015). http://ghtorrent.org/downloads.html.

[3] 2015. GitHub. (2015). https://github.com.

[4] Matt Asay. 2015. 3 key elements that define every open source. (2015). http://readwrite.com/2013/12/11/open-source-diversity.

[5] Olof Åslund and Oskar Nordströum Skans. 2012. Do anonymous job application procedures level the playing field? *Industrial & labor relations review* 65, 1 (2012), 82–107.

[6] Drake Baer. 2015. Psychologists say white men benefit from these unconscious biases. (2015). http://www.businessinsider.in/Psychologists-say-white-men-benefit-from-these-unconscious-biases/articleshow/46374544.cms.

[7] Marianne Bertrand and Sendhil Mullainathan. 2003. *Are Emily and Greg more employable than Lakisha and Jamal? A field experiment on labor market discrimination.* Technical Report. National Bureau of Economic Research.

[8] Christian Bird, Alex Gourley, Prem Devanbu, Anand Swaminathan, and Greta Hsu. 2007. Open borders? immigration in open source projects. In *Mining Software Repositories, 2007. ICSE Workshops MSR'07. Fourth International Workshop on.* IEEE, 6–6.

[9] Donn Erwin Byrne. 1971. *The attraction paradigm.* Vol. 11. Academic Pr.

[10] Jacob Cohen, Patricia Cohen, Stephen G West, and Leona S Aiken. 2013. *Applied multiple regression/correlation analysis for the behavioral sciences.* Routledge.

[11] Harald Cramér. 1999. *Mathematical methods of statistics.* Vol. 9. Princeton university press.

[12] John W Creswell. 2013. *Research design: Qualitative, quantitative, and mixed methods approaches.* Sage publications.

[13] Shane Curcuru. 2015. Open Source Is Way Whiter And Maler Than Proprietary Software. (2015). http://readwrite.com/2013/12/11/open-source-diversity.

[14] Elizabeth Dwoskin. 2015. Computers are showing their biases and tech firms are concerned. (2015). http://www.wsj.com/articles/computers-are-showing-their-biases-and-tech-firms-are-concerned-1440102894.

[15] Steve Easterbrook, Janice Singer, Margaret-Anne Storey, and Daniela Damian. 2008. Selecting empirical methods for software engineering research. In *Guide to advanced empirical software engineering.* Springer, 285–311.

[16] Uwe Flick. 2009. *An introduction to qualitative research.* Sage.

[17] Adrian Furnham. 1986. Response bias, social desirability and dissimulation. *Personality and individual differences* 7, 3 (1986), 385–400.

[18] Georgios Gousios, Martin Pinzger, and Arie van Deursen. 2014. An exploratory study of the pull-based software development model. In *Proceedings of the 36th International Conference on Software Engineering.* ACM, 345–355.

[19] Georgios Gousios, Margaret-Anne Storey, and Alberto Bacchelli. 2016. Work Practices and Challenges in Pull-Based Development: The Contributor's Perspective. In *Proceedings of the 38th International Conference on Software Engineering (ICSE).* DOI: http://dx.doi.org/10.1145/2884781.2884826

[20] Georgios Gousios and Andy Zaidman. 2014. A Dataset for Pull-based Development Research. In *Proceedings of the 11th Working Conference on Mining Software Repositories (MSR 2014).* ACM, New York, NY, USA, 368–371. DOI: http://dx.doi.org/10.1145/2597073.2597122

[21] Marek Hlavac. 2015. *stargazer: Well-Formatted Regression and Summary Statistics Tables.* http://CRAN.R-project.org/package=stargazer R package version 5.2.

[22] Sujin K Horwitz and Irwin B Horwitz. 2007. The effects of team diversity on team outcomes: A meta-analytic review of team demography. *Journal of management* 33, 6 (2007), 987–1015.

[23] Rachel L Johnson, Somnath Saha, Jose J Arbelaez, Mary Catherine Beach, and Lisa A Cooper. 2004. Racial and ethnic differences in patient perceptions of bias and cultural competence in health care. *Journal of general internal medicine* 19, 2 (2004), 101–110.

[24] Marie Kraska-Miller. 2013. *Nonparametric statistics for social and behavioral sciences.* CRC Press.

[25] Jennifer Marlow, Laura Dabbish, and Jim Herbsleb. 2013. Impression formation in online peer production: activity traces and personal profiles in github. In *Proceedings of the 2013 conference on Computer supported cooperative work.* ACM, 117–128.

[26] Nora McDonald and Sean Goggins. 2013. Performance and participation in open source software on github. In *CHI'13 Extended Abstracts on Human Factors in Computing Systems.* ACM, 139–144.

[27] Raphael Pham, Leif Singer, Olga Liskin, Fernando Figueira Filho, and Klaus Schneider. 2013. Creating a shared understanding of testing culture on a social coding site. In *Software Engineering (ICSE), 2013 35th International Conference on.* IEEE, 112–121.

[28] Emily Pronin, Daniel Y Lin, and Lee Ross. 2002. The bias blind spot: Perceptions of bias in self versus others. *Personality and Social Psychology Bulletin* 28, 3 (2002), 369–381.

[29] R Core Team. 2015. *R: A Language and Environment for Statistical Computing.* R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/

[30] Baishakhi Ray, Daryl Posnett, Vladimir Filkov, and Premkumar Devanbu. 2014. A large scale study of programming languages and code quality in github. In *Proceedings of the 22nd ACM SIGSOFT International Symposium on Foundations of Software Engineering.* ACM, 155–165.

[31] Anna Sandberg. 2015. Competing biases: Effects of gender and nationality in sports judging. (2015).

[32] Walt Scacchi. 2006. Free/open source software development: recent research results and methods. *ADVANCES IN COMPUTERS* 69 (2006), 244.

[33] Walt Scacchi. 2007. Free/open source software

development: Recent research results and methods. *Advances in Computers* 69 (2007), 243–295.

[34] Meir Shemla, Bertolt Meyer, Lindred Greer, and Karen A Jehn. 2014. A review of perceived diversity in teams: Does how members perceive their team's composition affect team processes and outcomes? *Journal of Organizational Behavior* (2014).

[35] LEE SIGELAMAN. 1981. Question-order effects on presidential popularity. *Public Opinion Quarterly* 45, 2 (1981), 199–207.

[36] Josh Terrell, Andrew Kofink, Justin Middleton, Clarissa Rainear, Emerson Murphy-Hill, and Chris Parnin. 2016. *Gender bias in open source: Pull request acceptance of women versus men*. Technical Report. PeerJ PrePrints.

[37] Jason Tsay, Laura Dabbish, and James Herbsleb. 2014. Influence of social and technical factors for evaluating contribution in GitHub. In *Proceedings of the 36th international conference on Software engineering*. ACM, 356–366.

[38] Bogdan Vasilescu, Vladimir Filkov, and Alexander Serebrenik. 2015a. Perceptions of diversity on GitHub: A user survey. *CHASE. IEEE* (2015).

[39] Bogdan Vasilescu, Daryl Posnett, Baishakhi Ray, Mark GJ van den Brand, Alexander Serebrenik, Premkumar Devanbu, and Vladimir Filkov. 2015b. Gender and tenure diversity in GitHub teams. In *CHI. ACM*.

[40] Quang H Vuong. 1989. Likelihood ratio tests for model selection and non-nested hypotheses. *Econometrica: Journal of the Econometric Society* (1989), 307–333.

[41] Peter Weißgerber, Daniel Neu, and Stephan Diehl. 2008. Small patches get in!. In *Proceedings of the 2008 international working conference on Mining software repositories*. ACM, 67–76.

[42] Joan C. Williams. 2015. The 5 biases pushing women out of stem. (2015). https://hbr.org/2015/03/the-5-biases-pushing-women-out-of-stem.