Technical Report

September 24, 2020

This is the appendix of TSE paper "pull-based development".

1 Results for decision making with different contexts

1.1 Decision making

1.1.1 The maturity of project

To find whether there is difference for the influence of factors on decision making for projects with different maturity. We split the data into three categories according to the time gap between project creation and data collection (project history in month), namely young projects ($proj\_hist \leq 29$), semi-mature projects ($29 < proj\_hist \leq 55$) and mature projects ($proj\_hist > 55$). The result can be seen in Table 1.

1.1.2 The popularity of project

In order to detect whether factors will have different influence on decision making for projects with different popularity. We split the data into three categories according to the star number of projects when collecting data, namely unpopular projects ($star\_num \leq 1952$), semi-popular projects ($1952 < star\_num \leq 35153$) and popular projects ($star\_num > 35153$). The result can be seen in Table 2.

1.1.3 The programming language of project

We also split the dataset according to the programming language of project. There are 6 programming languages, namely Java, Javascript, Python, Ruby, Go, Scala. The result of each programming language is shown in Table 3.

1.2 Latency

Similar to decision making, we also split the projects according to maturity (Section 1.1.1), popularity (Section 1.1.2) and program language (Section 1.1.3).

1.2.1 The maturity of project

The result can be seen in Table 4.

Table 1: Factors influence on decision making for projects with different maturity

	Dependent variable: merged_or_not=1		
	Young	Semi-mature	Mature
(Intercept)	18.412***	14.694***	12.085***
prior_review_num	1.883***	1.878***	1.654***
same_user	0.463***	0.514***	0.573***
lifetime_minutes	0.524***	0.591***	0.653***
core_member	1.569***	1.674***	1.484***
has_comments	0.674***	0.736***	0.631***
num_commits	1.311***	1.264***	1.239***
ci_exists	1.187***	1.135***	1.133***
open_pr_num	0.630***	0.615***	0.770***
watchers	0.815***	0.890***	1.016*
hash_tag	1.184***	1.135***	1.134***
project_age	0.891***	1.169***	1.068***
followers	1.197***	1.121***	1.184***
src_churn	1.163***	1.117***	1.005
team_size	1.000	1.012	1.102***
commits_on_files_touched_close	1.059***	1.011**	1.107***
files_changed	0.839***	0.853***	0.865***
test_inclusion	1.084***	1.114***	1.071***
Observations	467,705	1,002,719	1,699,797
Note:			*p<0.1; **p<0.05; ***p<0.01

Table 2: Factors influence on decision making for projects with different popularity

	Dependent variable: merged_or_not=1		
	Unpopular	Semi-popular	Popular
(Intercept)	21.974***	16.263***	9.697***
prior_review_num	1.730***	1.747***	1.711***
same_user	0.500***	0.484***	0.563***
lifetime_minutes	0.485***	0.551***	0.654***
core_member	1.498***	1.678***	1.522***
has_comments	0.680***	0.710***	0.653***
num_commits	1.266***	1.229***	1.257***
ci_exists	1.130***	1.163***	1.150***
open_pr_num	0.504***	0.663***	0.771***
stars	1.111**	1.000	1.052***
hash_tag	1.101***	1.099***	1.147***
project_age	1.111***	1.060***	1.000
followers	1.033**	1.118***	1.191***
src_churn	1.188***	1.128***	1.021***
team_size	0.970	1.009	1.070***
commits_on_files_touched	1.138***	1.079***	1.073***
files_changed	0.892***	0.869***	0.858***
test_inclusion	1.104***	1.132***	1.073***
Observations	270,144	549,751	2,256,377
Note:			*p<0.1; **p<0.05; ***p<0.01

Table 3: Factors influence on decision making for projects with different programming language

	Dependent variable: merged_or_not=1					
	Java	Python	JavaScript	Ruby	Go	Scala
(Intercept)	20.153***	18.347***	14.464***	19.819***	23.345***	13.026***
prior_review_num	1.638***	1.678***	1.769***	1.564***	2.282***	1.582***
same_user	0.521***	0.534***	0.569***	0.505***	0.568***	0.691***
lifetime_minutes	0.681***	0.628***	0.566***	0.634***	0.602***	0.728***
core_member	1.472***	1.586***	1.583***	1.601***	1.565***	1.272***
has_comments	0.674***	0.612***	0.677***	0.633***	0.766***	0.724***
num_commits	1.205***	1.211***	1.330***	1.232***	1.324***	1.170***
ci_exists	1.212***	1.155***	1.096***	1.064***	1.206***	1.091***
open_pr_num	0.794***	0.741***	0.637***	0.790***	0.913***	0.715***
stars	1.025	1.069***	0.900***	1.006	0.788***	1.081*
hash_tag	1.234***	1.116***	1.097***	1.148***	1.180***	1.222***
project_age	1.039***	0.987	1.149***	1.040***	0.873***	0.990
followers	1.123***	1.153***	1.168***	1.272***	1.183***	1.027**
src_churn	0.986**	1.058***	1.116***	1.021**	0.960***	1.076***
team_size	1.099***	1.094***	1.004	1.099***	1.142***	0.929***
commits_on_files_touched	1.119***	1.062***	1.023***	1.102***	1.067***	1.245***
files_changed	0.879***	0.876***	0.864***	0.846***	0.867***	0.834***
test_inclusion	1.046***	1.149***	1.071***	1.069***	1.046***	0.927**
Observations	625,688	878,660	935,042	343,558	269,479	117,794

Note: *p<0.1; **p<0.05; ***p<0.01

Table 4: Factors influence on latency for projects with different maturity

	Dependent variable: lifetime_minutes		
	Young	Semi-mature	Mature
(Intercept)	0.532***	0.658***	0.689***
description_length	1.139(12.2%)***	1.143(11.4%)***	1.162(10.7%)***
same_user0	1.428(20.2%)***	1.429(20.1%)***	1.460(19.7%)***
core_member1	0.877(2.7%)***	0.863(2.5%)***	0.853(2.2%)***
has_comments1	1.698(45.5%)***	1.669(43.9%)***	1.746(47.3%)***
num_commits	1.149(5.6%)***	1.133(7.6%)***	1.124(10.7%)***
ci_exists1	1.144(0.7%)***	1.112(0.6%)***	1.109(1%)***
open_pr_num	1.116(3.2%)***	1.220(3.1%)***	1.235(1.1%)***
watchers	0.983(0.5%)***	0.913(0.4%)***	0.898(0%)***
hash_tag1	1.185(2.8%)***	1.187(4%)***	1.167(3.5%)***
project_age	1.040(0%)***	1.052(0.4%)***	0.993(0.3%)***
followers	0.990(0.2%)***	0.987(0.1%)***	0.979(0%)***
src_churn	1.090(4.4%)***	1.115(4.6%)***	1.116(3.1%)***
team_size	0.983(0.7%)***	0.952(0.2%)***	0.921(0%)***
commits_on_files_touched	0.980(0.7%)***	0.968(0.6%)***	0.962(0.2%)***
files_changed	0.988(0.2%)***	0.982(0.1%)***	0.977(0.1%)***
test_inclusion1	1.050(0.4%)***	1.053(0.3%)***	1.063(0.3%)***
Observations	462,884	988,963	1,670,593
Note:			*p<0.1; **p<0.05; ***p<0.01

1.2.2 The popularity of project

The result can be seen in Table 5.

Table 5: Factors influence on latency for projects with different popularity

	Dependent variable: lifetime_minutes		
	Unpopularity	Semi-popular	Popular
(Intercept)	0.563***	0.665***	0.688***
description_length	1.163(11.8%)***	1.158(11.8%)***	1.151(12.5%)***
same_user0	1.493(19.5%)***	1.488(21.4%)***	1.431(22%)***
core_member1	0.839(2.4%)***	0.840(3.1%)***	0.866(3.4%)***
has_comments1	1.737(45.6%)***	1.723(45.5%)***	1.714(45%)***
num_commits	1.149(6.2%)***	1.139(7.7%)***	1.124(8.7%)***
ci_exists1	1.133(0.5%)***	1.112(0.7%)***	1.100(0.9%)***
open_pr_num	1.177(2.9%)***	1.188(2.3%)***	1.218(2.2%)***
star	0.889(0.5%)***	0.949(0.1%)***	0.903(0.3%)***
hash_tag1	1.139(3.7%)***	1.171(2.4%)***	1.179(1.4%)***
project_age	1.060(0%)***	1.046(0.3%)***	1.017(0.8%)***
followers	0.997(0.2%)	0.981(0.1%)***	0.979(0%)***
src_churn	1.083(4.9%)***	1.106(3.6%)***	1.119(2.2%)***
team_size	0.977(0.6%)***	0.956(0.2%)***	0.926(0%)***
commits_on_files_touched	0.966(0.7%)***	0.963(0.6%)***	0.965(0.5%)***
files_changed	0.997(0.2%)	0.982(0.1%)***	0.977(0%)***
test_inclusion1	1.049(0.4%)***	1.046(0.2%)***	1.063(0.2%)***
Observations	267,625	543,877	2,217,488
Note:			*p<0.1; **p<0.05; ***p<0.01

1.2.3 The programming language of project

The result can be seen in Table 6.

Table 6: Factors influence on latency for projects with different programming language

	Dependent variable: lifetime_minutes					
	Java	Python	JavaScript	Ruby	Go	Scala
(Intercept)	0.583***	0.619***	0.611***	0.582***	0.637***	0.658***
description_length	1.140(8.5%)***	1.139(10.4%)***	1.179(15.3%)***	1.149(11.6%)***	1.134(10.2%)***	1.113(9%)***
same_user0	1.457(21.5%)***	1.440(19.9%)***	1.525(22.5%)***	1.430(19.5%)***	1.329(14.5%)***	1.216(8.6%)***
core_member1	0.839(3.3%)***	0.872(2.2%)***	0.868(2%)***	0.829(4%)***	0.877(2.3%)***	0.845(4.5%)***
has_comments1	1.752(48%)***	1.707(45.5%)***	1.732(44.4%)***	1.687(44.6%)***	1.663(41.2%)***	1.642(41%)***
num_commits	1.104(4.6%)***	1.138(7.5%)***	1.126(5.7%)***	1.139(7.2%)***	1.164(11.3%)***	1.154(13.6%)***
ci_exists1	1.157(1.5%)***	1.061(0.2%)***	1.094(0.4%)***	1.172(1.4%)***	1.121(0.7%)***	1.166(1.9%)***
open_pr_num	1.221(3%)***	1.206(3.1%)***	1.224(2.9%)***	1.147(1.4%)***	1.142(1.3%)***	1.123(1.2%)***
stars	0.888(0.5%)***	0.886(0.6%)***	0.922(0.3%)***	0.901(0.4%)***	0.922(0.4%)***	0.941(0.2%)***
hash_tag1	1.196(3%)***	1.172(3.3%)***	1.146(2%)***	1.147(2.3%)***	1.209(6.7%)***	1.257(10%)***
project_age	1.016(0%)***	1.036(0.2%)***	1.036(0.2%)***	1.036(0.2%)***	1.093(1%)***	1.043(0.3%)***
followers	0.980(0.2%)***	0.987(0.1%)***	0.988(0.1%)***	0.986(0.1%)***	0.968(0.7%)***	0.948(2.1%)***
src_churn	1.100(4.1%)***	1.119(4.6%)***	1.100(3.1%)***	1.138(5.4%)***	1.143(7%)***	1.130(6.5%)***
team_size	0.959(0.2%)***	0.902(1.3%)***	0.953(0.2%)***	0.937(0.4%)***	0.989(0%)**	0.953(0.3%)***
commits_on_files_touched	0.955(1%)***	0.976(0.3%)***	0.966(0.6%)***	0.954(1.1%)***	0.962(0.9%)***	0.970(0.6%)***
files_changed	0.999(0%)	0.969(0.3%)***	0.975(0.2%)***	1.000(0%)	0.951(1%)***	0.986(0.1%)***
test_inclusion1	1.069(0.6%)***	1.075(0.5%)***	1.028(0.1%)***	1.065(0.4%)***	1.082(0.9%)***	0.910(0.1%)***
Observations	615,908	864,016	922,114	338,495	265,898	116,009

Note: *p<0.1; **p<0.05; ***p<0.01