Air Quality, Knowledge Worker Performance, and Adaptation: Evidence from GitHub*

Felix Holub

Department of Economics, Goethe University Frankfurt

Beate Thies

Department of Economics, University of Vienna

December 2023

Abstract

Highly skilled knowledge workers are important drivers of innovation and long-run growth. We study how air quality affects productivity and work patterns among these workers, using data from *GitHub*, the world's largest coding platform. We combine panel data on daily output, working hours, and task choices for a sample of more than 27,000 software developers across four continents with information on concentrations of fine particulate matter ($PM_{2.5}$). An increase in air pollution reduces output, measured by the number of total actions performed on GitHub per day, and induces developers to adapt by working on easier tasks and by ending work activity earlier. To compensate, they work more on weekends following high-pollution days, which suggests adverse impacts on their work-life balance. Exposure to unusually high $PM_{2.5}$ levels relative to the city-by-season-by-day-of-week specific mean reduces output quantity by 4.4%, which translates into a daily loss in output value by approximately $8.4 per developer.

JEL Codes: D24, J22, J24, L86, Q52, Q53

Keywords: Knowledge workers, Productivity, Adaptation, GitHub, Air pollution

---

*We thank Ulrich Wagner, Tatyana Deryugina, Robert Elliot, Johannes Gessner, Kathrine von Graevenitz, Andreas Grunewald, Nicolas Koch, Hannah Klauber, Simon Krause, Giovanni Paolo Mariani, Alessandro Palma, Nico Pestel, Linnéa Rohde, Konrad Stahl, Kinga Tchorzewska, Nicolas Ziebarth, and audience members at several conferences, workshops, and seminars. Support by the German Research Foundation (DFG) through CRC TR 224 (Project B07) is gratefully acknowledged. This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (Grant agreement No. 865181).

Corresponding author: Beate Thies, beate.thies@univie.ac.at

1 Introduction

Driven by technological innovation, the world of work is undergoing rapid changes. Over the last decades, computerization has been causing an increase in the demand for workers performing non-routine, analytical, and interpersonal tasks (Autor et al., 2003; Autor and Price, 2013). Skills that complement digital technologies have been growing in importance: In the US, the share of jobs requiring intensive digital skills, such as the ability to handle information technology or to conduct data analyses, has more than quadrupled, increasing from 5% to 23% between 2002 and 2016 (Muro et al., 2017). Work organization has also changed, with teamwork, flexible schedules, and task discretion replacing traditional 9-to-5 schedules and direct task assignments, particularly among highly-educated workers (Bresnahan et al., 2002; Mas and Pallais, 2020; Menon et al., 2020). Because jobs characterized by these task profiles, skill requirements, and organizational features form the backbone of the modern knowledge economy and are expected to become even more important as digitization and automation proceed, it is critical to understand what determines productivity in such settings.

In this paper, we study how environmental shocks impact performance and work patterns among highly skilled knowledge workers in a flexible work environment. Vast populations are exposed to environmental conditions such as heat and poor air quality, which have been shown to reduce labor productivity in several settings. Existing research, however, has considered routine jobs and/or inflexible work contexts (e.g. Chang et al., 2019; Somanathan et al., 2021). In the settings described above, workers not only employ different skills, but they also have flexibility and discretion in organizing their workday. This may enable them to adapt to productivity shocks, potentially alleviating output effects. Moreover, in collaborative work settings, impacts of environmental shocks might get dampened, e.g., if co-workers can help each other to focus, or get amplified due to complementarities if co-workers rely on each others' input.

We focus on the effects of air pollution, a ubiquitous public health concern in urban areas across the globe. Around 82% of the global population are exposed to levels of fine particulate matter that exceed World Health Organization (WHO) guidelines. We study the causal effect of air pollution exposure on professional software developers, using high-frequency data from *GitHub* to measure developer output and work patterns. As a STEM (science, technology, engineering, and math) occupation, software development requires analytical and advanced digital skills and generates high value for consumers, other industries, governments, and the research community.[1] Therefore, adverse productivity effects of air pollution in this occupation would have important implications for growth, innovation, and competitiveness. GitHub, the world's largest online code hosting platform, is used for storing and jointly working on coding projects and puts great emphasis on facilitating collaboration between developers. Moreover, software developers often work in flexible settings that offer discretion over working hours

[1] Median annual pay of software developers in the US was $110,140 in 2020 (Bureau of Labor Statistics, 2021).

and task selection. These features make software development on GitHub a representative setting for modern knowledge work.[2]

The GitHub data allow us to address the challenge that the output of knowledge workers is often difficult to observe. We collect data on 27,000 users who work on projects owned by tech companies, indicating that they are professional software developers. The data provide users' locations as well as records of all actions they conduct in public projects along with precise timestamps, including for instance commits (individual code changes) and comments written in discussion fora. We construct a user-by-day panel including measures of work quantity and quality, working hours, and task choice for the period between February 2014 and May 2019. Based on developers' locations, we match these outcomes to city-level air quality monitor data on particulate matter smaller than 2.5 µm ($PM_{2.5}$). To account for endogeneity in air quality, we follow previous research (e.g. Arceo-Gomez et al., 2016; Jans et al., 2018; Sager, 2019) by instrumenting $PM_{2.5}$ concentration with local temperature inversions. The instrumental variable (IV) strategy exploits the effect of plausibly exogenous changes in vertical temperature profiles in the atmosphere on local air pollution concentrations, controlling for a wide range of other weather characteristics.

Our dataset covers 47 countries, including both developing and developed countries, with large variation in pollution levels, income, and pollution awareness across sample cities. We exploit this in heterogeneity analyses to explore how air pollution damages are distributed and to study the mechanisms underlying the pollution impacts.

We present three main findings. First, developers produce less output on days with higher levels of fine particulate matter. When $PM_{2.5}$ concentration reaches unusually high levels of at least one standard deviation above the city-by-month-by-day of week-specific average, the number of daily actions observed on GitHub falls by 4.4%. We find substantial effect heterogeneity by type of work: The adverse effect of pollution is mainly driven by a decline in *individual* coding activity and work on new tasks. The number of commits for instance decreases by 5.9%. By contrast, *collaborative and responsive* work (e.g., commenting on issues) is less affected. Compared to other occupations studied in previous research, including both physically and cognitively demanding jobs, the magnitude of the effects on output is moderate. Nonetheless, the pollution-induced output declines translate into relevant monetary damages due to the high value generated by software developers. The loss in output value per developer and day amounts to \$3.63 for a standard deviation increase in $PM_{2.5}$ and \$8.4 for days with unusually high air pollution.

Second, output quality is only marginally affected by changes in air pollution. We find a minimal decrease in the acceptance rate of code changes a developer proposed on a given day when $PM_{2.5}$ increases. A potential explanation for the near-zero impacts on quality, as

[2]Evidence in Appendix Figure 1 and Table 1 shows that software development resembles other high-skilled occupations in skill requirements like critical thinking (except for a greater need for digital skills), and features like work flexibility and frequent teamwork.

well as the moderate size of effects on quantity, is worker adaptation to pollution-induced productivity shocks in this flexible work setting.

Our third main result provides evidence for this: On days with unusually high $PM_{2.5}$ levels developers work on easier tasks. Code they submit or review changes 3.2% fewer files and contains 4.1% fewer new lines compared to low pollution days, indicating that the code addresses less complex tasks. Among developers with stronger adjustment in task choice in response to $PM_{2.5}$ exposure, effects on output quantity and quality are attenuated. Software developers also adapt their working hours, shifting work from high-pollution, low-productivity weekdays to low-pollution, high-productivity weekends. Notably, developers end work activity earlier on days with unusually high $PM_{2.5}$ concentration. To compensate, they work more on weekends after a workweek with poor air quality, especially if air pollution on the weekend is moderate, compensating for 39% of the output reduction on the day of exposure. These adjustments likely explain why, compared to other professions, we find moderate effects of particulate matter on output. However, this shift to weekend work implies an additional welfare costs due to forgone leisure time and potential adverse impacts on work-life balance.

The adverse effects of pollution on output quantity arise at concentrations below current regulatory standards in the European Union and the US, with the strongest effects at lower levels of $PM_{2.5}$. The effect magnitude does not vary systematically with country-level pollution awareness, indicating that the negative effects on output are not driven by avoidance behavior. This is corroborated by the fact that we find relatively small extensive margin effects. Effects are substantially larger in locations with an older building stock, suggesting that differences in effective indoor pollution exposure play an important role. This points to a physiological mechanism underlying our main results. We also find that, within continents, lower-income cities experience stronger adverse effects.

To investigate whether pollution exposure also causes long-run consequences, we repeat our analysis at the monthly level. Air pollution exposure reduces the number of actions performed per month, with an effect size comparable to the analysis at the daily level. In addition, pollution decreases the monthly growth rate of a developer's number of followers on GitHub, a summary measure of work quantity, quality, and relevance, indicating that it not only reduces short-run performance but also slows down the build-up of reputation, which could plausibly have long-run consequences for developers' career paths.

Our paper makes several contributions to the literature. First, our work demonstrates new ways to use publicly available data on GitHub activity. While we are not the first to use this data in economics (for example, current work by McDermott and Hansen (2021) analyzes the impacts of the COVID-19 pandemic on GitHub work patterns), we propose strategies to construct a sample of highly active users who are likely professional software developers, to study task difficulty, and to estimate the monetary value of the output observed on GitHub.

We also directly link to the literature on air pollution and worker productivity. Several studies document a negative impact of pollution on productivity in manual and routine jobs, such as textile workers or call center agents (Adhvaryu et al., 2022; Chang et al., 2016, 2019; Graff Zivin and Neidell, 2012; He et al., 2019). A few papers also explore the effects of poor air quality on performance in more cognitively demanding work settings, including error detection by baseball umpires (Archsmith et al., 2018), quality of speeches held by politicians (Heyes et al., 2019), and case handling time by judges (Kahn and Li, 2020; Sarmiento, 2022). Related work investigates performance in cognitively demanding tasks outside of work settings, e.g., among chess or brain game players (Huang et al., 2020; Krebs and Luechinger, forthcoming; Künn et al., 2023; La Nauze and Severnini, 2021). While these settings provide precise performance measures in a specific domain, they do not capture the typical features of work organization in modern high-skill jobs. Furthermore, the rather inflexible settings studied so far do not allow to analyze worker adaptation to pollution. We expand the analysis to a STEM profession that is representative for a large group of knowledge workers in modern, flexible, and collaborative work environments. Thereby, our study adds novel insights into the labor market cost of air pollution in the context of ongoing digitalization. Also, we present first evidence on productivity effects separately for individual and collaborative activities, a distinction absent from previous work. Furthermore, our study, unlike others that often rely on data from a single country or location, uses a large, international sample of developers. This allows us to draw more general conclusions about the pollution-productivity relationship and to explore effect heterogeneity with respect to local income levels or pollution awareness.[3]

Finally, we contribute to the literature on worker adaptation to environmental shocks and connect it to research on the effect of flexible work arrangements on productivity. Several papers study how workers adjust working hours in response to extreme temperatures (Graff Zivin and Neidell, 2014; Neidell et al., 2021; LoPalo, 2023). Concerning air pollution shocks, Adhvaryu et al. (2022) and Bassi et al. (2021) demonstrate how managers mitigate productivity losses, for example by reallocating workers to different tasks. While these studies focus on rather low-skilled manufacturing workers, our work identifies new margins of adjustment in a flexible high-skilled setting, namely task choice and temporal reallocation of work activity towards the weekend.[4] By showing that workers exploit their flexibility to adapt to a productivity shock, and thereby alleviate its adverse impact, our work links to research on the causal effects of flexible work arrangements and worker autonomy on productivity (e.g. Beckmann et al., 2017; Angelici and Profeta, 2020). Our results suggest that the ability to adapt to idiosyncratic productivity shocks might contribute to the positive relationship between flexible work arrangements and performance.

[3] Borgschulte et al. (forthcoming) and Fu et al. (2021) analyze pollution impacts on US labor earnings and Chinese manufacturing productivity, without focusing on specific professions. We add new evidence to this with our international sample and our analysis of worker adaptation which requires high-frequency microdata.

[4] In parallel work on workers in Mexico City, Hoffmann and Rud (2022) find that workers reallocate labor supply across days to avoid pollution exposure. In contrast, we center on highly-skilled STEM workers and interpret the shift as compensating for reduced productivity rather than avoidance.

2 Background on Particulate Matter

We focus on $PM_{2.5}$, particulate matter with a diameter of less than 2.5 μm, which encompasses all solid and liquid particles suspended in the air. In urban areas, most $PM_{2.5}$ originates from anthropogenic sources, including traffic, industrial production, and biomass burning (Karagulian et al., 2015).

The main reasons to focus on $PM_{2.5}$ are (i) its ability to penetrate indoors, making it particularly relevant for indoor office workers (Deng et al., 2017), and (ii) the fact that a large body of research documents that fine particulate matter plays a key role for the adverse health effects of air pollution. The small particles can penetrate deep into the lungs, causing damage to the respiratory system, as well as cardiovascular health effects like high blood pressure and heart diseases (Lederer et al., 2021). While severe health effects are concentrated among vulnerable groups, even healthy adults can experience mild symptoms, including irritation in the nose and throat or coughing (Pope, 2000). In response to the evidence on adverse health effects, several countries introduced standards on annual ambient $PM_{2.5}$ concentrations and often tightened them over time. Currently, standards are in place for example in the US (12 $\mu g/m^3$) and the European Union (25 $\mu g/m^3$). The WHO recommends a level of no more than 5 $\mu g/m^3$.

Recent clinical and autopsy studies suggest that exposure to fine particles can even affect the central nervous system as the small particles can reach the brain, causing neuro-inflammation (Babadjouni et al., 2017). In line with this, higher $PM_{2.5}$ concentrations have been found to lower scores in online brain games (La Nauze and Severnini, 2021) and high-stakes exams (Ebenstein et al., 2016).

Motivated by the prior research on particulate matter and its impact on health and cognitive functioning, we seek to *quantify* productivity impacts among knowledge workers in a modern work environment and investigate potential adaptation responses in such settings.

3 Setting and Data

This section starts with a brief description of GitHub, followed by an overview of the GitHub data and how we use it to measure developers' productivity. After checking the validity of these outcome measures, we describe the air quality and weather data.

3.1 Git and GitHub

GitHub is built on *Git*, an open-source version control system that tracks changes made to files, recording who made the change and when. *GitHub* is a web platform for hosting Git repositories, the directories where all files belonging to specific projects are stored. On top of the version control functionality, GitHub also provides collaboration tools. For each repository (or repo for short), the owner can choose whether to make it public (visible to everyone) or private (visible only to repository members). In 2019, GitHub had more than 30 million registered accounts owning more than 120 million public repositories, making it the world's largest host of source code.

The core action in Git is a *commit*, which refers to saving the current state of the repository after making changes to one or more files. As such, a commit represents that some work on code files was conducted by the commit author. Only repository owners and invited team members can modify files through commits.

The primary collaboration features offered on GitHub are *pull requests* and *issues*. A pull request (PR) is used to propose code changes to a repo. To create a PR, a user creates a copy of the repository, implements changes in their copy via commits, and then submits these to the original repository. Repo members then review the proposed changes and decide whether to accept (i.e., merge) or reject them. Each PR includes a discussion forum where users can comment directly on the proposed changes. Feedback provided here can be implemented within the same PR. Hence, PRs facilitate collaborative coding and are used not only for external contributions but also within project teams.

Issues are text messages typically used to organize tasks within a repo. Like PRs, issues contain a discussion forum where users can comment on the matter at hand. Repository members can assign labels to issues to highlight their category (e.g., bug, feature request), priority, or difficulty. GitHub provides nine default labels, and teams can create additional labels specific to their repo. Once an issue is resolved, it can be marked as closed.

Additionally, GitHub has social network functions, e.g., options to follow other users and subscribe to specific repos and issues to receive notifications about new activities.

Since GitHub actions related to commits, PRs, and issues reflect productive work aimed at building or improving software products, we collect data on these activities to measure output generated by highly skilled developers.

3.2 GitHub Data on Productivity and Work Patterns

The *GHTorrent* project provides a database on GitHub users and all their actions in *public* repositories. We use the version of the database containing data up to June 1st, 2019. The *user* table comprises a unique identifier, login name and registration date for all users registered by that date. In addition, location and company information as stated on the user profile on this date is reported. The *projects* table provides identifiers and names of all public repositories as well as a reference to the user owning the repo. Activity data is available by type (e.g., commits, opening issues, PR comments), including exact timestamps and the identifiers of the acting user and the repository where the event occurred. For certain actions, further information is reported, such as labels attached to issues.

We complement this with data from *GHArchive*, which also records actions in public repositories and contains additional information on some events, e.g., the title of a commit (commit message), the number of lines of code added and deleted, and the number of files changed within a PR. GHArchive and GHTorrent data can be linked via users' login names.

These data have several favorable features for our analysis. First, the precise records of GitHub activities allow us to quantify the daily output of software developers, addressing the long-standing challenge of measuring the work of highly skilled workers during a given period. Second, the data cover all GitHub users, providing a broader geographic coverage and thus an advantage in terms of external validity compared to previous studies. Moreover, the rich information included allows us to measure not only output quantity but also quality and work patterns, which are crucial in knowledge work.

The data also have limitations. To assign local air quality to users, we rely on self-reported locations, which may be wrong or outdated, giving rise to measurement error. This, under the assumption of classical measurement error, leads to attenuation bias such that any adverse effect found represents a lower bound. Additionally, we lack information on work conducted in private repositories or outside GitHub. Many users engage in no or minimal work in public repositories, making it impossible to measure their productivity based on their activity data. Thus, when constructing our analysis sample, we aim to capture users who are professional software developers and do a substantial part of their formal work in public GitHub repositories.

Sample Construction. We focus on users who report a location at the city level so that we can assign local air quality. From this group, we only keep users who have ever committed in a repository owned by a company, i.e., users with the authority to change the source code of a company-owned project. This step is intended to focus on professionals who are affiliated with the companies. To identify these users, we compile a list of company-operated repositories[5] and then use the information on the repository where a commit was made from the GHTorrent data. To exclude bots—computer programs often used to automate routine tasks—we discard a small number of users with bot-like behavior.[6]

To focus on cases where we observe a substantial part of an individual's total work, we admit users to the sample once they made at least 20 commits in public repos in a given month. They enter the sample in the month after meeting this threshold for the first time. Users remain in the sample until the end of the observation period unless they conduct fewer than three *unproductive* actions in a given month, which include following other users, watching repositories, (un)subscribing to issues, labeling issues, and (un)assigning issues to users. In such cases, we drop users from the sample for that month, assuming they moved to a different platform or private repos. Unproductive actions are not used as outcomes and are based

[5] This list is based on https://github.com/d2s/companies/blob/master/src/index.md and lists of open-source projects operated by Google, Microsoft and Facebook mentioned on their websites.

[6] We exclude users in the top 0.1 activity percentile, with over 20% of commits occurring on exact full hours, or with bot-indicative login names.

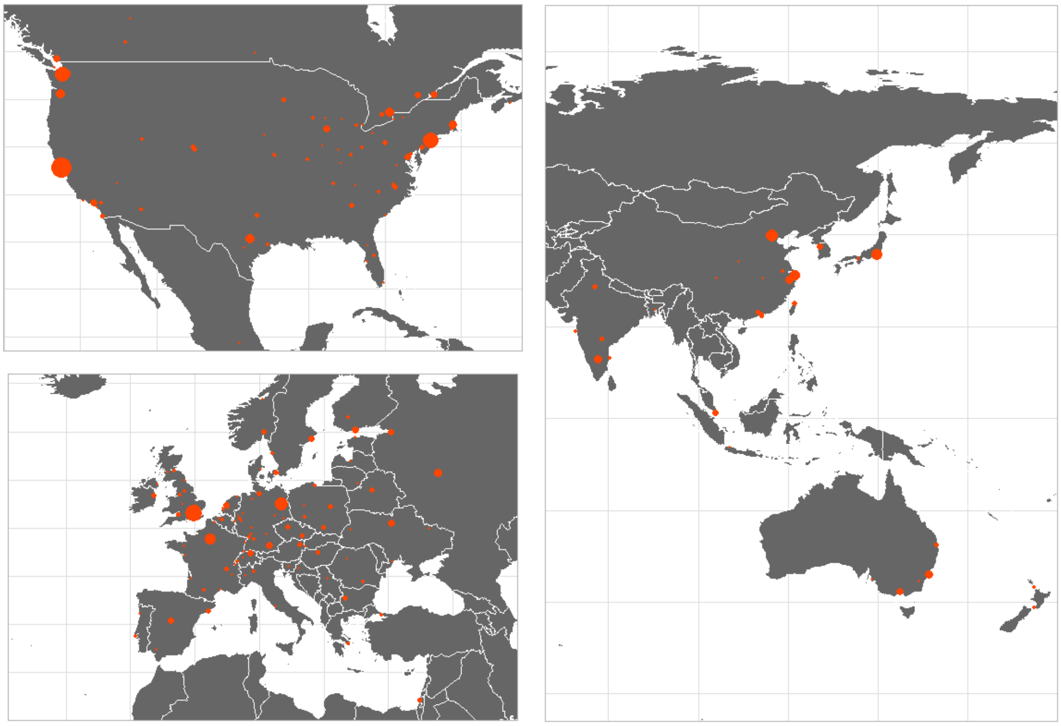


Figure 1: Sample Cities
Note: Points represent sample cities. Circle size is based on the number of users observed in the city.

primarily on the social network functions GitHub offers. Lastly, we restrict the sample to users living in cities with at least 15 relevant users and covered by our air pollution data. This yields 27,686 users across 220 cities in 47 countries (Figure 1).

Outcome Measures. We compile an unbalanced user-by-day panel for the period between February 2014 and May 2019, including measures of output quantity, quality, and work patterns.[7] To measure overall daily work *quantity*, we count the total number of productive actions conducted per user and day, given as the sum of commits, PR and issue creations, PR and issue closings, issue reopenings, and comments on PRs, issues or commits. Furthermore, we separately count commits, as a measure of individual coding work, and comments, as a measure of interactive activity in response to work conducted by others, to account for potential effect heterogeneity between these two distinct activity categories.

To assess output *quality*, we measure (i) the percentage of PRs opened on a given day that are merged, indicating acceptance, and thus satisfactory code quality, and (ii) the share of commits made on a given day that get reverted at a later point, signaling significant issues with the code that cannot be easily rectified.[8]

[7] During our sample period some users changed their location. Since the GHTorrent data on users is a snapshot taken on June 1st, 2019, we use earlier versions of the database (one snapshot in each year between 2015 and 2018) to check for movements. 6.3% of users reported more than one distinct location. We identify the city where they spent the biggest part of the sample period, and keep them in the sample only while residing in this city.

[8] Revert commits have specific commit messages allowing to identify them and the commit that is reverted.

To assess worker *adaptation*, we build measures of task choice and working hours. Firstly, to explore whether users switch to easier tasks on high-pollution days, we consider the complexity of issues and PRs. For PRs, we count the number of new lines of code added, lines of code deleted, and code files changed. We take the average value of these variables across all PRs a user worked on a given day, either by creating or reviewing the respective PR. For issue difficulty, we rely on the user-assigned issue labels which indicate that a given issue is relatively easy, e.g., the default labels *good first issue* and *documentation* (Tan et al., 2020), or individual labels such as *low-hanging fruit*. The complete list of labels we use to identify easy tasks is depicted in Appendix Table 2. We construct the share of all issue events conducted by the user (commenting, opening, closing, or reopening of issues) which refer to an *easy* issue. With this approach, we do not have to evaluate issue complexity ourselves but rely on the experts' assessment. Furthermore, the label is visible to all users, i.e., workers searching for easy tasks can easily identify suitable issues. Secondly, we investigate whether users try to make up for a productivity shock by working longer hours in the evening or on the weekend. Evening work is measured by the time of the last action of the day (in minutes since midnight). Weekend work is the sum of actions conducted on Saturdays and Sundays.

Finally, as a summary measure of the quantity, quality, and relevance of a user's work, we consider the monthly growth rate of their followers. This allows to investigate whether air pollution shocks have only temporary effects, or also impair users' reputation and influence over a more extended period.

All outcomes, with details on their construction, are listed in Appendix Table 3.

Descriptives. While our sample comprises just 0.085% of all GitHub users, they conduct a disproportionally large share of activities in public repositories, namely 2% of issue creations, 8% of issue closings, 11% of comments, and 7% of PR openings and closings. Hence, the selected users are a highly active subsample. This is confirmed by the summary statistics on the outcome variables (Table 1). The average number of daily actions is 2.68, of which 1.27 are commits and 0.88 are comments. The remaining productive GitHub actions—opening and closing issues and PRs—occur less frequently. On average, users are active on 36% of all days during the sample period, and conditional on being active, the mean number of actions is 7.45. These figures imply high activity levels, especially given that they include weekends and holidays. Commit reversals, which signify significant errors, are rare, occurring in only 0.2% of all commits. The share of PRs not getting accepted is higher at 33%. On average, 7% of all issue events (opening, closing, reopening, or commenting on an issue) refer to an easy issue. The mean time of the final action of the day is 5:32 pm.[9]

Figure 2 depicts the distribution of activity across days of the week and hours of the day. The solid lines represent the share of all activity conducted during each hour of the day on weekdays (left) or weekends (right) for commits, comments, and total actions. Activity levels

[9]To account for the tendency of high-skill workers to work long hours in the evening, we define a workday from 3 am on the calendar date to 3 am on the following day.

Table 1: Summary Statistics for the Analysis Sample of GitHub Users

	Mean	SD	SD (within)	Min	Max	Observations
Output Quantity						
Actions	2.68	7.09	6.31	0	293	16,499,962
of which Commits	1.27	3.78	3.51	0	234	16,499,962
Comments	0.88	3.26	2.85	0	280	16,499,962
PRs opened	0.15	0.69	0.65	0	151	16,499,962
Issues opened	0.10	0.79	0.73	0	222	16,499,962
PRs closed	0.16	0.92	0.87	0	284	16,499,962
Issues closed	0.11	0.87	0.85	0	263	16,499,962
Any action	0.36	0.48	0.43	0	1	16,499,962
Actions | Actions > 0	7.45	10.22	9.08	1	293	5,925,638
Output Quality						
PR Acceptance Rate	0.67	0.45	0.40	0	1.0	1,237,267
Share commits reverted	0.002	0.027	0.027	0	1	4,448,241
Task Complexity						
Share easy issue events	0.07	0.21	0.20	0	1.0	3,515,595
Files changed per PR	6.68	18.57	17.97	0	641	1,781,745
Lines added per PR	259.24	1271.62	1239.28	0	41681	1,781,745
Lines deleted per PR	104.80	594.43	582.11	0	19778	1,781,745
Working Hours						
Time last action	17:32	5.03 hours	4.66 hours	3:00	3:00	5,902,740

Notes: Table shows key measures of output quantity, quality, complexity, and working hours at the developer×date level.

peak during core working hours and decline in the evening and night hours and on weekends. Notable activity during evening hours and on weekends is not uncommon among highly educated workers (Mas and Pallais, 2020). The distribution is similar across all three variables, but comments, i.e., more interactive activities, are even more concentrated during standard working hours compared to commits, i.e., individual coding activities. This is plausible given that more collaborative tasks are more productive at times when other users are working as well.

Finally, Figure 3 presents information about the work status of users in our sample. The left plot depicts the most frequent terms used in the biographies (bios) on their GitHub profiles. 35% of the users provide such a self-description. We measure the occurrence of each term in the bios after stemming and removing stop words. Three terms stand out: *engineer/engineering*, *software*, and *developer/development*. The right plot complements this with information on employers reported by users on their profiles. 65% of the users provide some information in this field. The most frequently reported employers are major tech companies strongly engaged in open-source. While we cannot assess if the users who provide this information are representative for the whole sample, the high prevalence of work-related terms and notable tech firms, along with the peak in activity during core working hours, strongly indicates that

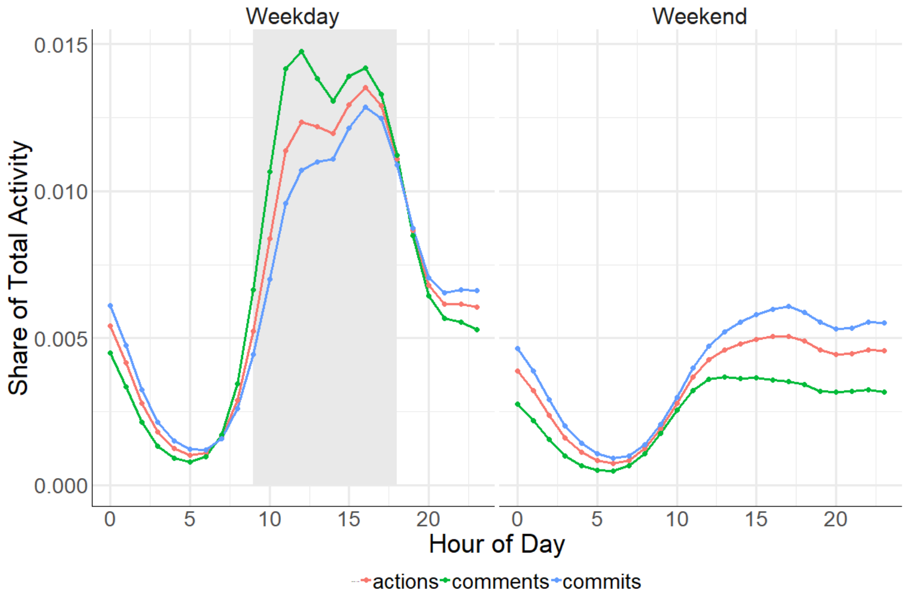


Figure 2: Distribution of Activity across Hours of the Day and Days of the Week

Note: Hourly share of activities on weekdays (left) or weekends (right), for total actions, comments, and commits. Grey area: core working hours, 9 am to 6 pm on weekdays.

we are capturing professional developers using GitHub for formal work. Thus, in the remainder of the paper, we refer to the sample users as 'developers'.

3.3 Gitcoin: Monetary Value of GitHub Activity

To validate our performance metrics and to translate the effects of air pollution into monetary damages, we draw on data from the web platform *Gitcoin*. On Gitcoin, GitHub project teams aiming to incentivize external contributions to their projects post issues from their public repos and announce a payment they offer for a solution. Freelance developers can apply to solve these issues and earn the payments for their contributions. Work on the issues is submitted in the form of a PR in the respective GitHub repo. If the PR is accepted by the issue funder, the PR author receives the payment, typically in cryptocurrencies.

We collected data on 292 issues for which PRs were submitted and payments made by March 2022 via the Gitcoin API, including the value of the payment in USD and the hours worked on the PR as reported by the submitting user. We merge this with information on the size of the respective pull request obtained via the GitHub API (number of commits, number of lines of code added and deleted, and number of files changed). A detailed description of the data is provided in Appendix C.

We find mean payments of $354 per pull request and $112 per commit. Developers spend on average 1.8 hours per commit. This implies an hourly wage of $62, very close to the mean wage of $58 among software developers in the US in 2021 (Bureau of Labor Statistics, 2021).

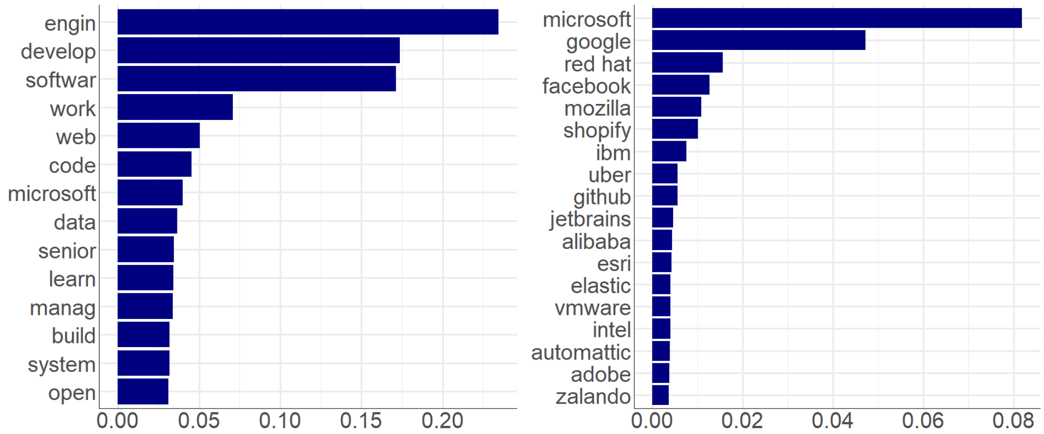


Figure 3: Most Frequent Terms from User Self-Descriptions and Company Fields

Note: Left: Most frequent terms from user self-descriptions, after stemming and removing stopwords. Data from 9,696 user bios, accessed in 2021 via the GitHub API. Right: Most frequent companies. Data from June 2019 on 17,928 users. Horizontal axes measure share of users reporting the respective term or company.

We will use the monetary values of commits and PRs to translate the effects of air pollution into monetary damages.

Are the outcomes we consider valid measures of productivity and task complexity? In Appendix Tables 1 to 3 we use the Gitcoin data to test this. Both, payments awarded through Gitcoin for PRs and the time spent on PRs, are positively correlated with the number of commits that PRs comprise. This confirms that changes in the number of commits reflect fluctuations in developer productivity. Additionally, holding the number of commits constant, adding more lines of code and changing more files in a PR is associated with a higher payment, suggesting that these variables indeed reflect task complexity. 12% of the Gitcoin issues are labeled as *easy* according to our definition. Even when controlling for all previously mentioned PR characteristics, these issue labels are associated with a 35% decline in the payment for a PR, supporting their validity as indicators for easy tasks.

3.4 Environmental Data

Air Quality. We collect $PM_{2.5}$ concentration data from outdoor monitors from several environmental agencies. For cities without publicly available monitor data, we use high-resolution reanalysis data from the Copernicus Atmosphere Monitoring Service (CAMS). Reanalysis data are constructed by combining ground-level measurements, satellite images, and atmospheric transport models. Appendix Table 4 provides a list of the data sources. The data are provided at daily or hourly intervals. We adjust hourly data to local time before aggregating to the daily level. Cities are assigned the simple average of all monitor readings within a 40 km radius around the city centroid. For cities covered by the CAMS data, we only use points within 25 km due to the high resolution. Our data on $PM_{2.5}$ covers 95% of observations in the GitHub panel.

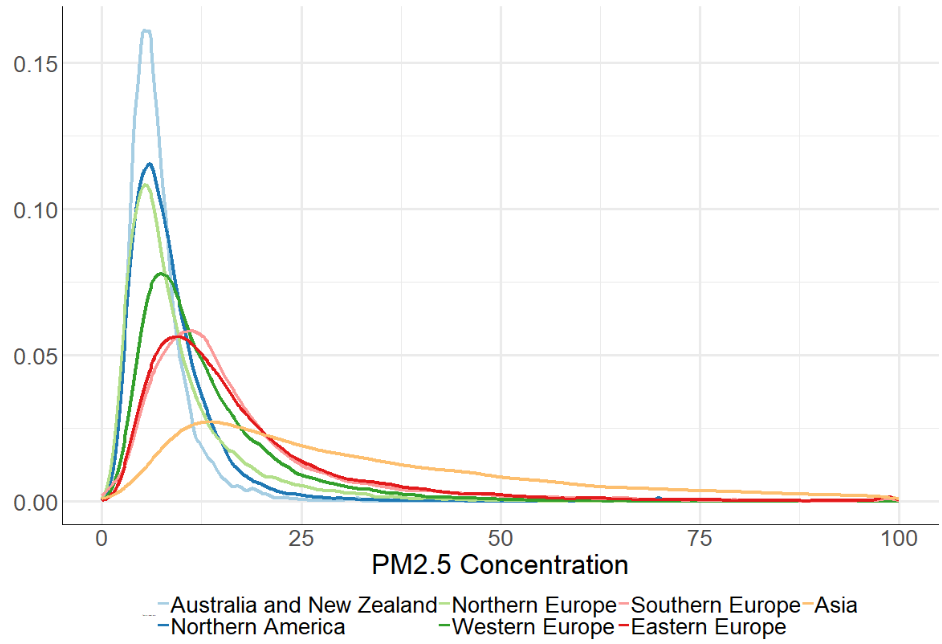


Figure 4: Regional distribution of daily $PM_{2.5}$

Notes: Oceania: Australia, New Zealand. Northern Europe: Scandinavia, UK, Ireland, Baltics. Southern Europe: Portugal, Spain, Italy, Greece, Israel, Croatia, Serbia, Slovenia, Turkey. Asia: Bangladesh, China, India, Indonesia, Japan, Korea, Hong Kong, Singapore, Taiwan. Northern America: US, Canada, Mexico. Western Europe: Switzerland, Austria, France, Germany, Belgium, Netherlands. Eastern Europe: Poland, Czech Republic, Hungary, Belarus, Ukraine, Slovakia, Bulgaria, Romania, Russia.

We winsorize $PM_{2.5}$ at the continent-specific 0.1th and 99.9th percentile to mitigate the influence of extreme outliers, such as heavy wildfire smoke or industrial accidents. The population-weighted average $PM_{2.5}$ concentration in the sample is 15.15 μg/m^3 (standard deviation: 22.3 μg/m^3, within-city: 17.6 μg/m^3), which is between the regulatory standards on annual $PM_{2.5}$ concentration in the US (12 μg/m^3) and the EU (25 μg/m^3). Figure 4 displays the distribution of daily $PM_{2.5}$ for seven large geographic regions, $R \in$ {Northern Europe, Southern Europe, Western Europe, Eastern Europe, North America, Oceania, Asia} and demonstrates substantial regional heterogeneity. While concentrations above 20 μg/m^3 are infrequent in North America, Oceania, and Northern Europe, cities in Southern and Eastern Europe experience this level of pollution on 26% of all days, and Asian cities as much as 65% of the time.

Thermal Inversions. The instrumental variable approach is based on thermal inversions, which are defined as atmospheric conditions where the temperature of upper air layers is *higher* than the temperature at the earth's surface. Under normal conditions, air temperature decreases with altitude. To construct inversion measures, we collect reanalysis data on hourly temperature at the surface level and at several pressure levels from the European Centre for Medium-Range Weather Forecasts (ECMWF) ERA5 products. The data is reported on a 0.25-degree grid, corresponding to roughly 28 km × 28 km at the equator. For each grid point and hour, we compute the difference between upper air temperature and surface air temperature in degrees Celsius. Upper air temperature is measured at the pressure level 25 hPa above the surface pressure level, which corresponds to an altitude difference of approximately 210 meters.

We then calculate the average temperature difference during local nighttime hours (midnight to 6 am), after translating timestamps into local time. To obtain city × day level variables, each city is assigned the inverse distance weighted average nighttime temperature difference from the four grid points closest to its centroid. This continuous measure of inversion strength will be our main instrument for local pollution. In robustness checks, we use an inversion indicator as an alternative instrument, which takes value one if the difference between upper air and surface temperature is positive and zero otherwise. Inversions occur on 46% of all days during the sample period, and the median temperature difference is -0.3°C, with the first and third quartiles at -1.7°C and 2.1°C, respectively. Appendix Table 5 shows the frequency of inversions by geographic region.

Further Meteorological Conditions. To construct control variables for daily weather conditions we use the ERA5-land product from the ECMWF. It provides hourly data on air temperature, precipitation, and dewpoint temperature on a 0.1-degree grid. We compute daily mean, minimum and maximum temperature, precipitation, and relative humidity at each grid point and then assign each city the inverse distance weighted average weather conditions from the eight closest grid points.

We collect reanalysis data on wind conditions from the Japan Meteorological Agency's JRA-55 product. The u- and v-component of wind, i.e., the wind vectors, are reported every six hours on a 1.25-degree grid. We aggregate to the daily level and compute the inverse distance weighted average of u- and v-vectors at the four grid points located closest to each city's centroid. Finally, daily average wind speed and direction are computed from the city-level u- and v-vectors.

The North American west coast, where some of the largest cities in our sample are located, frequently experiences severe wildfires, leading to major peaks in air pollution due to smoke. Recent research has shown that wildfire smoke exposure can trigger avoidance behavior (Burke et al., 2022). To ensure that our estimates capture physiological impacts of $PM_{2.5}$ exposure rather than behavioral responses to wildfires, we construct control variables for heavy smoke events. The National Oceanographic and Atmospheric Administration's Office of Satellite and Product Operations provides data on the location and intensity of smoke plumes across North America. We define a city as being affected by wildfire smoke if a plume overlaps with a 15 km radius around its centroid, and aggregate the data to the daily level by summing over the intensities of all plumes covering a city. We define a heavy smoke indicator taking value one if the city was covered by a plume of the highest intensity or if the sum exceeds twice the maximum intensity. 0.3% of all city-day observations and 9% of all observations with any smoke exposure are classified as heavy smoke days.

4 Research Design

Baseline Regression Model. To analyze how short-run variation in local particulate matter concentration affects output and work patterns of professional software developers, we specify a model for developers living in city c on day d:

$$y_{c,d} = \beta PM_{c,d} + \mathbf{w}'_{c,d}\gamma_{R(c)} + \delta_{R(c)}h_{c,d} + \mu_c + \mu_{R(c),dow(d)} + \mu_{R(c),yr(d),m(d)} + \varepsilon_{c,d}. \tag{1}$$

Here, $y_{c,d}$ denotes one of the measures of output quantity, quality, or work patterns described in the previous section. We obtain this variable through an auxiliary regression that controls for developer experience in using GitHub and a developer fixed effect. By doing so, we can reduce the computational burden without losing variation in the regressor of interest which is observed at the city-day level. This procedure is commonly employed (e.g. Currie et al., 2015) and asymptotically equivalent to estimating the underlying individual-level regressions (Donald and Lang, 2007). Appendix D provides a more detailed description.

$PM_{c,d}$ is a measure of particulate pollution and varies across cities c and days d. The fixed effect μ_c controls for time-invariant unobserved factors at the city level. Region-specific year × month fixed effects $\mu_{R(c),yr(d),m(d)}$ capture time-varying productivity shocks common to all developers in a geographic region R. Region-specific day-of-week fixed effects $\mu_{R(c),dow(d)}$ and an indicator for holidays, $h_{c,d}$, control for fluctuations in work patterns and productivity across days of the week and public holidays. These fluctuations are allowed to vary in intensity across different world regions. The vector $\mathbf{w}_{c,d}$ contains weather variables that can be correlated with air quality, inversions, and work patterns. It includes a series of indicator variables for daily mean temperature falling into bins defined based on the 5th, 10th, 20th, 35th, 65th, 80th, 90th, and 95th percentiles of the city-specific temperature distributions, cubic polynomials of precipitation, relative humidity, and wind speed, as well as a dummy indicating heavy wildfire smoke. The effects of all weather variables are allowed to differ across regions R, taking into account that for instance individuals in warmer regions might respond differently to unusually warm temperatures than individuals in colder regions. We weight all regressions by the number of underlying developer observations in each city–day cell and cluster standard errors at the city level.

While we include a wide range of controls to account for sorting into different cities or fluctuations in economic conditions, $PM_{c,d}$ may still be endogenous in Equation (1) due to unobservable factors that co-vary with particulate matter and productivity. Variations in local economic conditions can for instance affect air pollution and developers' output at the same time. Similarly, local events like a football match or the closing of a bridge may impact both traffic and work patterns. Besides omitted variable bias, a second issue is measurement error in individual pollution exposure, which we have to proxy for by city-level averages, generating attenuation bias in the OLS estimator.

IV Estimation. To address endogeneity, we follow several recent studies (e.g. Arceo-Gomez et al., 2016) and instrument local pollution levels with temperature inversions. During an inversion period, air temperature increases with altitude, such that a warm upper air layer acts as a ceiling, preventing pollution emitted at the surface level from dispersing and trapping it near the ground. Consequently, the same amount of surface-level emissions results in higher pollution concentrations compared to normal periods. This effect on air quality is stronger, the stronger the inversion, i.e., the larger the difference between the upper and surface air temperature. Our instrument for pollution will therefore be the temperature difference between upper and lower layers, $Inv\ Strength_{cd} = \Delta T_{cd}$.

Following Jans et al. (2018) and Sager (2019), we focus on nighttime inversion strength, as daytime inversions may be visually noticeable in some regions which could lead to behavioral adjustments. Nighttime inversions are less likely to trigger such responses, strengthening the exogeneity of the instrument.

Inversions can be correlated with weather conditions. For instance, calm winds tend to reduce the mixing of cold and warm air, creating favorable conditions for the formation of inversions. Weather conditions could also affect labor-leisure trade-offs (Graff Zivin and Neidell, 2014) and thereby the output of developers via channels other than air quality. Moreover, in many places, inversions exhibit seasonality, with higher frequency during winter. Therefore, it is important to control for the wide range of weather conditions in $\mathbf{w}_{c,d}$ and to include fixed effects capturing seasonality effects.

In our global sample, the impact of a temperature inversion on air quality is not uniform across all cities. When a warm upper air layer hinders pollution from dispersing, this causes a strong increase in pollution concentration in locations with high emissions, whereas areas with low levels of ground-level emissions are much less affected (Krebs and Luechinger, forthcoming). Additionally, the effects of inversions can differ depending on geographic and topographic factors, such as when mountains further impede air circulation. To account for this, we allow the impact of inversion strength on $PM_{2.5}$ to vary between groups of cities that are geographically close to each other and exhibit similar pollution levels. We use a k-means clustering algorithm to assign each city into a group g based on longitude, latitude, and average $PM_{2.5}$ concentration during the sample period. In our baseline specification, we form 25 groups, which are depicted in Appendix Figure 2

The first stage of the IV estimation is as follows, where the coefficient π^g is allowed to vary across city-groups $g \in \{1, 2, \ldots, 25\}$:

$$PM_{c,d} = \pi^g \Delta T_{cd} + \mathbf{w}'_{c,d}\gamma_{R(c)} + \delta_{R(c)} h_{c,d} + \mu_c + \mu_{R(c),dow(d)} + \mu_{R(c),yr(d),m(d)} + \varepsilon_{c,d}. \qquad (2)$$

Figure 5 illustrates the first stage relationship for five city groups. It shows binned scatter plots of residualized $PM_{2.5}$ concentration, after taking out the weather controls and fixed effects mentioned above, against inversion strength. A linear function fits the relationship

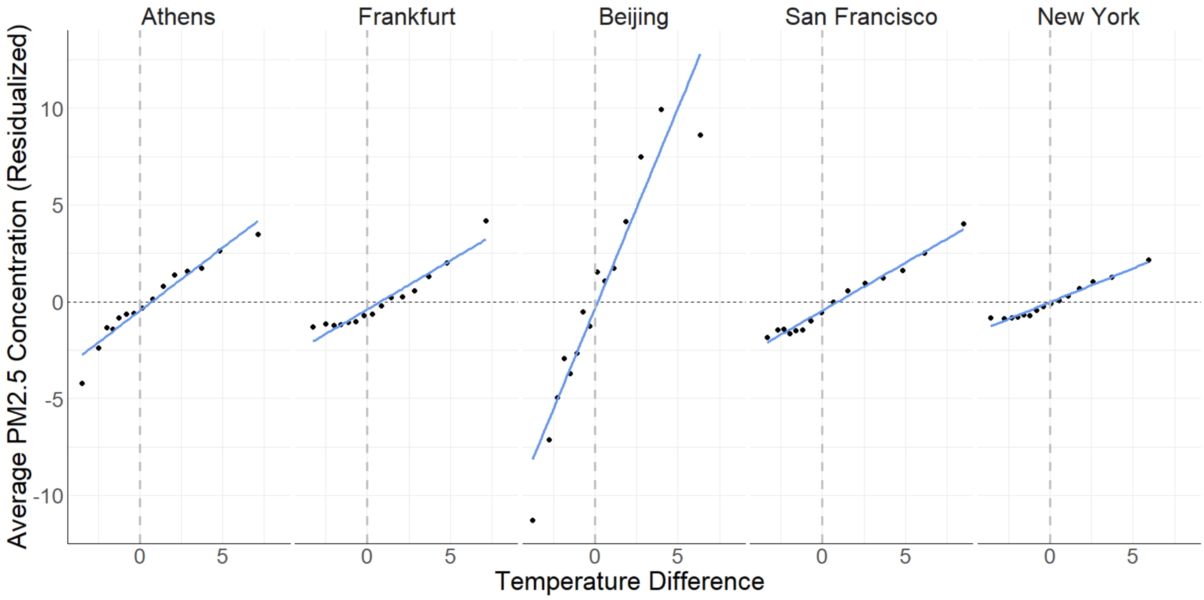


Figure 5: Illustration of the First Stage

Note: Binned scatter plots of residualized $PM_{2.5}$ and inversion strength in five city groups. Blue lines: linear fit. Groups: 9 Southeast European cities (*Athens*), 17 Central European cities (*Frankfurt*), 8 Chinese cities (*Beijing*), 16 cities in/around Silicon Valley (*San Francisco*), 15 cities in the Eastern US and Canada (*New York*).

between inversion strength and air pollution reasonably well. As anticipated, an increase in the temperature difference between upper and surface level air has a more pronounced effect on $PM_{2.5}$ concentration in cities with higher baseline pollution levels (e.g., Beijing vs. New York). In Appendix Figure 3 we present these plots for all 25 city groups.

Pollution Measures Our primary measure of air pollution is the daily average $PM_{2.5}$ concentration in $\mu g/m^3$. As an alternative measure, we define a binary variable that we will refer to as a *pollution shock*,

$$PM_{2.5}\, shock_{cd} \equiv \mathbf{1}\left\{PM_{c,d} > \sqrt{\widehat{\text{Var}}\ [PM \mid c]} + \widehat{\text{E}}\ [PM \mid c, m(d), dow(c)]\right\}. \tag{3}$$

The indicator takes value one if the city-day concentration of $PM_{2.5}$ is more than one city-specific standard deviation above the level expected for the given city, calendar month, and day of week. In essence, it reflects unusually high pollution levels in a location × season × day-of-week cell. This measure captures non-linear effects of pollution and allows these to vary by location and time. 10.6% of all city × day observations are characterized by such a $PM_{2.5}$ shock.

Table 2: IV Estimates of the Effect of $PM_{2.5}$ on Work Quantity

	Actions		*Commits*		*Comments*		*Any actions*	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$PM_{2.5}$	−0.0030		−0.0015		−0.0008		−0.0002	
	(0.0006)		(0.0004)		(0.0002)		(0.00004)	
	{<0.0001}		{0.0008}		{0.0011}		{<0.0001}	
$PM_{2.5}$ shock		−0.118		−0.075		−0.030		−0.010
		(0.066)		(0.029)		(0.030)		(0.005)
		{0.088}		{0.018}		{0.318}		{0.047}
Observations	397,277		397,277		397,277		397,277	
Mean dep. var.	2.68		1.27		0.88		0.36	
First stage F-stat.	160	76	160	76	160	76	160	76
% change	-0.11	-4.4	-0.12	-5.9	-0.09	-3.4	-0.06	-2.8
% of full effect							17.8	23.2

Notes: Regressor of interest is $PM_{2.5}$ in $\mu g/m^3$ in odd-numbered columns, or a dummy as in Equation (3) in even-numbered columns. Covariates: Eight bins for mean daily temperature, third-order polynomials in wind speed, precipitation, and relative humidity, indicators for heavy wildfire smoke and holidays, as well as city, region-by-day-of-week, and region-by-year-by-month fixed effects. Coefficients on weather controls can vary across regions. Regressions weighted by number of developers in city-month cell. Standard errors, clustered by city, in parentheses. Benjamini-Hochberg p-values in curly brackets.

5 Main Results

In this section, we first present results on how $PM_{2.5}$ exposure affects the quantity and quality of output developers produce. Thereafter we show that they use two margins of adjustment, task choice and working hours, to adapt to increases in pollution concentration.

5.1 Work Quantity

Table 2 displays 2SLS estimates of the effect of $PM_{2.5}$ exposure on output quantity. We find that an increase by 1 $\mu g/m^3$ causes developers' overall output, measured by total actions, to fall by 0.0030 or 0.11% of the sample mean. Half of this decline is driven by a reduction in the number of commits, our primary measure of individual coding activity, which decreases by 0.0015 or 0.12%. The decline in the number of comments written in discussion fora, as a measure of collaborative work, implies a slightly smaller reduction of 0.09% of the sample mean. The first stage F-statistic on the excluded instruments is 160, indicating that the IVs are sufficiently strong. For an increase in $PM_{2.5}$ by one within-city standard deviation (17.6 $\mu g/m^3$), the estimates imply reductions of 1.7% to 2.1% across the three outcomes. When we use the binary $PM_{2.5}$ shock variable as regressor, the F-statistic is again well above the common threshold for a sufficiently strong first stage. The 2SLS estimates imply that on a day with unusually high pollution, the number of total actions falls by 0.118 or 4.4% of the mean. The number of commits falls by 0.075 or 5.9%. The effect on comments is smaller and not statistically significant. In sum, these results imply that fine particulate matter exposure exerts a negative effect on developer output which is mostly driven by days with relatively poor air quality.

In Columns 7 and 8 we explore the contribution of the extensive margin to the overall reduction in work quantity. The dependent variable is an indicator for a positive activity level, i.e., $\mathbf{1}\{actions_{id} > 0\}$. For both measures of pollution, estimates are negative and significant, implying that the extensive margin effect contributes approximately 18-23% to the full reduction in actions. Hence, the intensive margin response is quantitatively more important. This result is plausible given that our sample of GitHub users likely comprises mostly young and middle-aged adults. They are unlikely to suffer severe health damages from short-run pollution exposure which prevent them from working, but rather experience subtle effects on health and cognitive function.

As we derived our results by testing eight hypotheses, we report p-values corrected for multiple hypothesis testing following the Benjamini-Hochberg procedure in curly brackets.

In Appendix Table 6, we investigate less frequent action types—the number of issues and PRs opened and closed, respectively. Like a commit, opening a PR reflects individual coding work. Closing a PR implies decision-making about whether to accept or reject proposed code changes, and opening/closing issues generally starts/ends a discussion with other users. Consistent with the above results, we find negative impacts of a 1 $\mu g/m^3$ increase in $PM_{2.5}$. The effects are largest for the number of new issues and PRs opened with -0.23% and -0.11% of the mean, respectively, compared to reductions by -0.08% in the number of PRs and issues closed. The pattern for $PM_{2.5}$ shocks is similar, but only the effect on opening new issues is statistically significant. Together with the results in Table 2, this implies that in response to higher air pollution, developers mostly reduce individual coding activity and activities on new tasks (commits, opening new issues), whereas effects are smaller for activities that respond to work conducted by others (writing comments, closing issues and PRs). This suggests that workers in collaborative work environments prioritize activities that involve co-workers when hit by an adverse productivity shock.

Columns 1 to 3 of Table 7 show OLS estimates for total actions, commits, and comments. We obtain negative estimates for all outcomes with both $PM_{2.5}$ in levels and the $PM_{2.5}$ shock indicator, but these are significant only for $PM_{2.5}$ in levels. Mirroring a common finding in the literature on air pollution impacts, all estimates are substantially smaller than the 2SLS results, pointing towards attenuation bias due to measurement error.

Effect Magnitude. We conduct three exercises to assess the magnitude and economic relevance of the estimates. Firstly, we compare the impact of a $PM_{2.5}$ shock to the effect of another highly relevant environmental shock, exposure to extreme outdoor temperatures, which have been found to affect, e.g., student performance, Twitter user sentiment, and mental health (Park, 2020; Baylis, 2020; Mullins and White, 2019). Figure 6 reproduces the estimated effects of a $PM_{2.5}$ shock on total actions and commits in graphical form (point estimates with confidence intervals displayed in black). In addition, coefficients from OLS regressions of the same outcomes on maximum daily temperature are presented. Specifically, we use eight dummy variables indicating whether maximum daily temperature falls into a certain percentile range of the city-specific distribution, as displayed on the x-axis. The reference category is

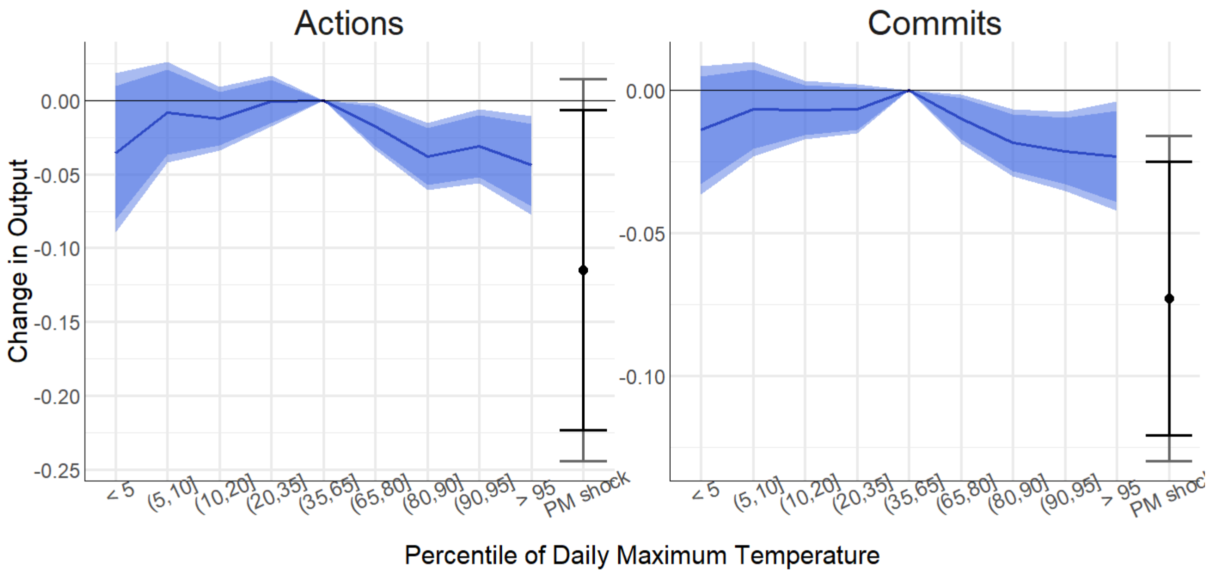


Figure 6: Effects of $PM_{2.5}$ and heat on work quantity

Note: Black points and bands: IV estimates with 90 and 95% confidence intervals for $PM_{2.5}$ shocks from Table 2. Blue lines and shades: OLS estimates with 90 and 95% confidence intervals for maximum daily temperature range indicators shown on the x-axis. Covariates: Flexible controls for minimum temperature, precipitation, wind speed and relative humidity, indicators for heavy wildfire smoke and holidays, as well as city, region-by-day-of-week, and region-by-year-by-month fixed effects. Standard errors clustered by city. Regressions weighted by number of developers in city-month cell.

maximum temperature between the 35th and the 65th percentile. For both outcomes, the effects of temperature are u-shaped: Both unusually cold and unusually hot temperatures have adverse effects, but only the impact of heat is statistically significant. Even though software developers might work in climate-controlled offices, heat exposure during commuting times or while running errands could generate these negative effects. While the IV estimates for $PM_{2.5}$ shocks, which occur on 10% of all days, are less precise than the OLS estimates for temperature, their point estimates are more than twice the estimates for the temperature bin representing maximum temperature above the 95th percentile. Hence, the adverse productivity effects of poor air quality exceed those of extreme temperatures, an environmental shock of high relevance given global warming.

Secondly, we compute elasticities for the effect of $PM_{2.5}$ on commits and total actions, and compare these to previous studies on productivity or performance effects. Our elasticity estimates of -0.017 and -0.018 are at the lower end of the range of effect sizes shown in Figure 7. The effect on developers' output is much smaller than estimates obtained for judges and chess players, who also engage in cognitively demanding tasks. A potential explanation is that chess tournaments and court hearings are highly inflexible settings that do not allow for adapting working hours or the choice of tasks to productivity shocks. This contrasts with our setting, and we provide evidence on worker adjustment to an increase in $PM_{2.5}$ in Section 5.3. This underscores the importance of our analysis: It might be misleading to estimate the total economic cost of air pollution based on inflexible settings if most knowledge workers have at least some degree of flexibility in organizing their workday.

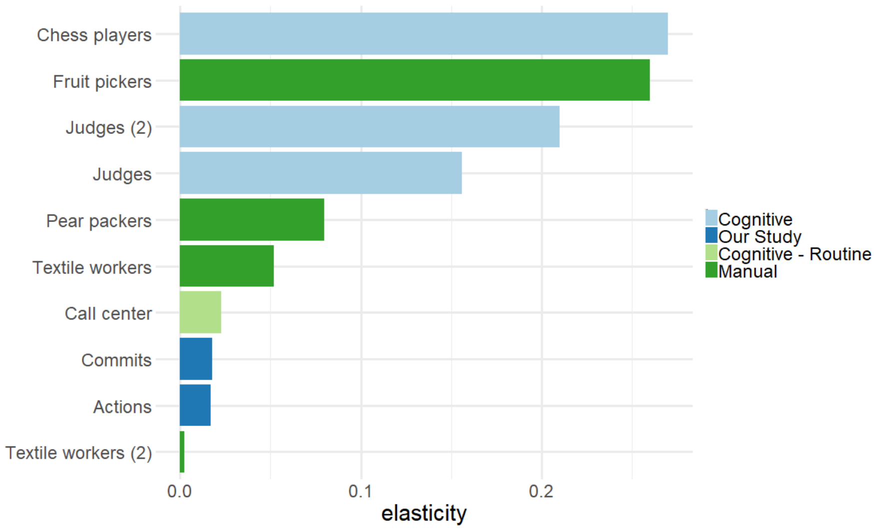


Figure 7: Effects of air pollution across occupations

Note: Elasticities of commits and actions based on Table 2. Elasticities from other studies: Künn et al. (2023) (chess players), Sarmiento (2022) (judges), Kahn and Li (2020) (judges (2)), Chang et al. (2019) (call center agents), He et al. (2019) (textile workers (2)), Adhvaryu et al. (2022) (textile workers), Chang et al. (2016) (pear packers), and Graff Zivin and Neidell (2012) (fruit pickers).

Even though productivity effects are small compared to other contexts, they might still be economically relevant, given that software development is an occupation generating large economic value. We use the average monetary value of commits and PRs opened (derived in Section 3.3) to translate the estimated effects of $PM_{2.5}$ exposure into changes in output value.[10] For a $10\,\mu g/m^3$ increase in $PM_{2.5}$ the implied reduction in daily output value amounts to $2.12 per software developer. This magnitude is comparable to effects reported by Chang et al. (2016) who find that a $10\,\mu g/m^3$ increase in $PM_{2.5}$ reduces hourly output among pear packers by $0.41, implying a damage of $3.28 for an eight-hour workday. On days with a $PM_{2.5}$ shock, output value falls by $8.40 relative to days with better air quality. As we ignore other losses, for example, from less work on issues and reductions in task complexity in the calculation, these estimates can be interpreted as lower bounds.

In summary, in comparison to other professions, the effect of particulate matter is moderate, pointing towards an important role of worker adaption in flexible work environments. Economically, the effect is nevertheless relevant, given the high monetary value of software.

Effect Dynamics. To explore effect dynamics, we use a "reduced form" version of our model, regressing output quantity on four variables measuring inversion strength in city *c* on day *d* and each of the previous three days. Appendix Table 8 shows that for all outcomes (total actions, commits, comments, and the indicator for any actions), same-day inversion strength generates negative and significant effects, whereas the point estimates on the lags are mostly close to zero and statistically not significant, except for a positive point estimate on total actions at the third lag. Thus, in our sample, which likely comprises mostly young to middle-aged,

[10] We assume a value of $112 per commit (Section 3.3). For PRs, we do not use the mean value of $354 found in the Gitcoin data because PRs in that sample are on average larger than PRs created in our sample. Instead, we value PRs with $2.13 \times \$112 = \$239$ given that they comprise, on average, 2.13 commits.

Table 3: IV Estimates of the Effect of $PM_{2.5}$ on Work Quality

	PR Acceptance Rate		*Share Commits Reverted*	
	(1)	(2)	(3)	(4)
$PM_{2.5}$	−0.00017		0.000001	
	(0.00009)		(0.000001)	
$PM_{2.5}$ shock		−0.0107		−0.0001
		(0.0079)		(0.0002)
First stage F-stat.	186	82	291	132
% change	−0.03	−1.6	0.03	-4.5
Observations	153,877		312,747	
Mean dep. var.	0.666		0.002	

Notes: Regressor is $PM_{2.5}$ in $\mu g/m^3$ in Colums (1) and (3) or a dummy as in Equation (3) in Columns (2) and (4). Covariates as in Table 2. Regressions weighted by number of developers in city-month cell. Standard errors, clustered by city, in parentheses.

high-skilled workers, pollution's adverse effects arise immediately and do not persist beyond the same day.

5.2 Work Quality

Apart from quantity, output quality is of major relevance, especially in high-skill jobs, and might also be affected by environmental shocks. Table 3 displays 2SLS estimates for two measures of work quality. The first is the share of PRs a user opened that later got accepted. PR rejections suggest issues with code quality or style, indicating low work quality. The second is the share of commits made by a user that were later reverted. Commit reversals point to severe errors that cannot easily be corrected in further commits, i.e., major issues with the work quality. Sample sizes are reduced, because these outcomes are only defined for city $\times$ day observations with any PRs opened and any commits, respectively. Moreover, information on commit reversals is from GHArchive, which is only available from 2015.

We find negative point estimates for the PR acceptance rate, implying a reduction by 0.017 percentage points or 0.03% relative to the sample mean rate for a one unit increase in $PM_{2.5}$, and a reduction by 1.07 percentage points or 1.6% on a day with a pollution shock relative to days with better air quality. The effect is statistically significant only when using $PM_{2.5}$ in $\mu g/m^3$ as regressor. Coefficients are insignificant for the share of commits that are reverted in both specifications. This might arise due to a lack of power to detect small impacts on this measure of severe quality issues.

Overall, we find only very small effects of pollution exposure on output quality. This contrasts with the results by Archsmith et al. (2018) who find quality effects that are almost an order of magnitude larger. They show that the propensity of baseball umpires to conduct errors increases by 2.6% in response to a 10 $\mu g/m^3$ increase in $PM_{2.5}$. In the next section, we present evidence that developers change their work patterns when exposed to high levels of pollution. This form of adaptation might explain why effects on output quantity are relatively moderate and quality is almost unaffected in this flexible high-skilled setting.

5.3 Worker Adjustment

To investigate whether work patterns change in response to increases in air pollution we consider two potential margins of adjustment: task choice and working hours.

Task Choice. We start by analyzing whether developers switch to easier tasks when exposed to higher levels of pollution. To assess the complexity of activities related to pull requests, i.e., coding and code review tasks, we consider three PR characteristics: lines of code added, lines of code deleted, and number of files changed, averaged across all PRs a developer worked on a given day.[11] While there can be very difficult tasks that involve a lot of effort but require only small changes in the code, we believe that these variables provide reasonable proxies of PR complexity. Fixing a severe bug likely requires changes in different parts of the code, implying a larger number of files changed. Results in Table 3 indicate that PRs with more lines of code added and files changed are rewarded higher payments on Gitcoin, validating the use of these variables as complexity metrics. Similarly, reviewing a PR is more demanding when it contains large changes across different files.

We apply the inverse hyperbolic sine transformation to the three variables such that coefficients approximate percentage changes. Table 4 presents the results. The sample size is reduced as the outcomes are defined only for city $\times$ day observations with any PRs opened or closed. An increase in $PM_{2.5}$ by 1 $\mu g/m^3$ reduces the three measures of PR complexity significantly, with effect sizes between -0.11% and -0.21% (Panel A). On days with a $PM_{2.5}$ shock, we find a 3.2% reduction in the number of files changed, and a marginally significant drop in the number of lines added by 4.1%, indicating that developers move towards less complex tasks mostly in response to large productivity shocks on high-pollution days (Panel B). The effect on the number of lines deleted is also negative, but small. This pattern is plausible as tasks related to deleting code, e.g., cleaning a file or dropping a redundant part, are usually easier to do and to review than tasks related to creating new code.

As an alternative to using the inverse hyperbolic sine transformation of the three PR characteristics, we summarise them in a single PR complexity index. It is computed by standardizing the mean number of lines added, of lines deleted, and of files changed per PR and taking the average across the three standardized measures. The resulting index is then divided by its standard deviation. Results for this index are reported in Table 10 and confirm that developers respond to increases in pollution by switching to easier PRs.

[11] These variables are based on GHArchive data, while work quantity results use GHTorrent data. GHArchive data are only available from 2015 onward. Appendix Table 9 shows that using data on PRs from GHArchive and GHTorrent yields very similar results for the effect of $PM_{2.5}$ on the number of PRs opened or closed.

Table 4: IV Estimates of the Effect of $PM_{2.5}$ on Task Complexity and Working Hours

	Lines added per PR (1)	*Lines deleted per PR* (2)	*Files changed per PR* (3)	*Share Easy Issue Events* (4)	*Time of Last Action (minutes)* (5)
Panel A.					
$PM_{2.5}$	−0.0021	−0.0017	−0.0011	0.00005	-0.0825
	(0.0008)	(0.0005)	(0.0003)	(0.00002)	(0.0355)
First stage F-stat.	200	200	200	260	195
% change	-0.2	-0.2	-0.1	0.07	
Panel B.					
$PM_{2.5}$ shock	−0.0413	−0.0107	−0.0317	0.0001	-5.149
	(0.0253)	(0.0252)	(0.0133)	(0.0024)	(2.499)
First stage F-stat.	87	87	87	117	67
% change	-4.1	-1.1	-3.2	0.15	
Observations	183,749	183,749	183,749	280,895	339,872

Notes: Panel A: Regressor is $PM_{2.5}$ in $\mu g/m^3$. Panel B: Regressor is a dummy as in Equation (3). In Columns (1) to (3), outcomes are transformed by the inverse hyperbolic sine. Covariates: See Section 4 or Table 2. Regressions weighted by number of developers in city-month cell. Standard errors, clustered by city, in parentheses.

Next, we analyze whether developers also focus on easier issue-related tasks. In Column (4), the outcome is the share of issue events completed that refer to an easy issue.We find an increase in the share in response to an increase in $PM_{2.5}$ concentration, but the magnitude is very small, and the result is not corroborated with the $PM_{2.5}$ shock measure. Even though certain issue labels provide a prominent signal of task complexity, developers do not seem to exploit this when hit by a pollution-induced productivity shock. A potential explanation is that less than 7% of all issues are labeled as “easy” so developers might not be able to find any open issues with such a label.

In sum, on top of the overall reduction in the number of actions completed, developers switch towards less complex coding and review tasks when exposed to high levels of $PM_{2.5}$. Thus, the estimates of monetary effects of air pollution exposure based on the reduction in output quantity provide a lower bound. This form of adjustment might also explain why the magnitude of effects on output quantity is moderate in comparison to results found in other settings, and why work quality is hardly affected.

To investigate this, we compare the effects of pollution on output between subsamples of developers who show a strong vs. weak adjustment response. We focus on developers who submitted or reviewed PRs on at least 100 days. For each developer, we run a separate IV regression of the PR complexity index on $PM_{2.5}$, the standard weather covariates, and time fixed effects. We store the estimated coefficients on $PM_{2.5}$, and split the sample into individuals with above or below median effect magnitude, i.e. weak or strong switch towards easier PRs in response to air pollution. We then run our IV regression for output quantity and quality at the developer $\times$ day level separately on the resulting subsamples.

Table 5: Effect Heterogeneity in Pollution Effects on Adjustment and Output

	Actions	*PR Acceptance Rate*	*Lines added per PR*	*Files changed per PR*
	(1)	(2)	(3)	(4)
Panel A. Strong Adjustment Response				
$PM_{2.5}$	-0.0082	-0.0002	-0.0080	-0.0045
	(0.0018)	(0.00016)	(0.0022)	(0.0014)
Observations	2,875,041	424,790	635,440	635,440
Mean dep. var.	5.14	0.68		
Panel B. Weak Adjustment Response				
$PM_{2.5}$	-0.0113	-0.0004	0.0035	0.0020
	(0.0044)	(0.0001)	(0.0028)	(0.0013)
Observations	2,870,976	429,163	637,461	637,461
Mean dep. var.	5.16	0.68		

Notes: Regressions are estimated at the developer × date level using two-stage least squares. In Columns (3) to (4), outcomes are transformed by the inverse hyperbolic sine. The samples used in Panel A and Panel B are from developers with more than 100 days with any PR activity, regressing the PR complexity index on $PM_{2.5}$ and covariates separately for each developer and then grouping developers by below vs. above median coefficient on $PM_{2.5}$, i.e. strong vs. weak adjustment response. Covariates: developer, region-by-day-of-week, and region-by-year-by-month fixed effects, bin variables for developers' experience on GitHub as well as weather covariates as described in Section 4. Standard errors, clustered by city, in parentheses.

The first two columns of Table 5 display estimated effects of $PM_{2.5}$ on the primary measures of work quantity (number of total actions) and quality (PR acceptance rate). Columns (3) to (4) show the effects on the adjustment measures. By construction, we find strong reductions in PR complexity in the *strong adjustment* subsample, and insignificant, smaller point estimates in the *weak adjustment* subsample. Point estimates for total actions and the PR acceptance rate are negative in both samples, but larger in absolute as well as relative terms for the developers showing no adjustment response. These results confirm that switching to easier tasks is a form of adjustment to pollution-induced productivity shocks among highly-skilled workers.

Working Hours. A second adjustment margin available is a change in working hours. We start by analyzing whether developers expand or reduce evening activity in response to pollution exposure. Column (5) in Table 4 presents the estimated effects of $PM_{2.5}$ on the timestamp of the last action performed by a developer on a given day (in minutes). We find that an increase in $PM_{2.5}$ concentration causes developers to end their work day slightly earlier, but the effect magnitude is very small. A $PM_{2.5}$ shock induces a stronger response, causing developers to end the work day on average 5 minutes earlier than on days with better air quality. In sum, developers do not use the evening to compensate for the productivity shock. Subtle effects of pollution might make them feel unproductive, inducing them to end their work activity earlier due to the lower opportunity cost of leisure. If $PM_{2.5}$ exposure triggers for example headaches or fatigue, developers might experience this as an off day and decide to reallocate work to days when they perform better.

In many jobs, knowledge workers are flexible in when and where they want to work. Shifting work intertemporally from low productivity days to the weekend, a period with relatively low activity levels and thus scope for compensation (see Figure 2), might therefore be an important adjustment margin. To investigate this, we estimate the effect of $PM_{2.5}$ exposure during the workweek on output produced on the weekend. This analysis is conducted at the developer × week level, using the following regression model:

$$
\begin{aligned}
y_{i,c,w}^{weekend} = \quad & \beta PM_{c,w}^{Mo-Fr} + \mu_i + \mathbf{x}'_{i,t}\pi + \mathbf{w}'^{weekend}_{c,w}\gamma_{R(c)} + \mathbf{w}'^{Mo-Fr}_{c,w}\alpha_{R(c)} + \\
& \delta_{R(c)} h_{c,w} + \mu_{R(c),yr(w),q(w)} + \mathbf{z}_{c,w}^{weekend}\varphi + \varepsilon_{i,c,w}.
\end{aligned} \tag{4}
$$

Here, $y_{i,c,w}^{weekend}$ denotes the sum of actions conducted by developer i living in city c on the weekend of week w. $PM_{c,w}^{Mo-Fr}$ measures the number of days with a $PM_{2.5}$ shock between Monday and Friday of week w. Due to the finding that exposure to pollution reduces output only on the same day without adverse effects on the following days (Appendix Table 8), the coefficient β should pick up developers' *behavioral* adjustment to a productivity shock during the workweek, and not confound it with *physiological* effects.

Pollution is instrumented by average nighttime inversion strength between Monday and Friday, interacted with indicators for the first-stage city groups g. To account for auto-correlation in the instruments, we add the vector $\mathbf{z}_{c,w}^{weekend}$ to the model, which includes the instrumental variables measured on the weekend. This ensures that we do not pick up the effects of inversion-induced changes in pollution on the weekend itself. The model further includes a developer fixed effect μ_i, a region-by-year-by-quarter fixed effect, $\mu_{R(c),yr(w),q(w)}$, the number of public holidays during the workweek, $h_{c,w}$, a vector $\mathbf{x}'_{i,t}$ of bin variables capturing the developer's tenure on GitHub, and two sets of weather controls, $\mathbf{w}_{c,w}^{Mo-Fr}$ and $\mathbf{w}_{c,w}^{weekend}$, covering the exposure period and the weekend, respectively.[12]

Results in Panel A of Table 6 indicate that developers produce significantly more output on weekends if they were exposed to unusually high levels of $PM_{2.5}$ during the workweek. One additional day with a $PM_{2.5}$ shock causes an increase in total actions on the weekend by 0.046 or 1.5% of the mean. Effects are positive for both commits and comments, but significant only for the latter, indicating an increase by 2.4% relative to the mean value. We also find positive point estimates for the probability of conducting any action at all (Column 4), and the PR complexity index, computed across all PRs opened or closed on the weekend (Column 5), but they are not statistically significant.

In Panel B we repeat the same analysis, but only based on weeks with low pollution levels on the weekend, defined as levels below the city-specific 75th percentile on both days. We find substantially larger coefficients in this sample, which attain statistical significance across all five outcomes. For one additional $PM_{2.5}$ shock during the workweek, developers are 0.4 percentage points more likely to conduct any work on GitHub on the weekend. The number of total actions increases by 0.11, driven by both more commits and more comments,

[12] The vectors $\mathbf{w}_{c,w}^{Mo-Fr}$ and $\mathbf{w}_{c,w}^{weekend}$ contain the number of days with wildfire smoke exposure, third-order polynomials in average precipitation, relative humidity, and wind speed during the respective period. The vector $\mathbf{w}_{c,w}^{Mo-Fr}$ further includes variables counting the number of days on which daily mean temperature falls into the temperature bins described above. The vector $\mathbf{w}_{c,w}^{weekend}$ includes a third-order polynomial in average temperature.

Table 6: Effect of $PM_{2.5}$ during Workweek on Weekend Work (IV Estimates)

	Actions (1)	*Commits* (2)	*Comments* (3)	*Any Actions* (4)	*PR Complexity Index* (5)
Panel A. Full sample					
Days with $PM_{2.5}$ shock	0.0457	0.0094	0.0176	0.0011	0.0172
	(0.0269)	(0.0143)	(0.0087)	(0.0012)	(0.0109)
First stage F-stat.	73.5	73.5	73.5	64.5	
Observations	2,273,906	2,273,906	2,273,906	2,273,906	207,967
% change	1.5	0.6	2.4	0.3	1.7
Panel B. Low PM Weekends only					
Days with $PM_{2.5}$ shock	0.1138	0.0437	0.0361	0.0038	0.0288
	(0.0501)	(0.0242)	(0.0158)	(0.0017)	(0.0122)
First stage F-stat.	51.0	51.0	51.0	51.0	58.7
Observations	1,518,897	1,518,897	1,518,897	1,518,897	138,933
% change	3.9	2.6	4.9	1.1	2.8

Notes: 2SLS regressions are estimated at the developer × week level. The regressor is the number of days with a $PM_{2.5}$ shock between Monday and Friday. Panel A: Models estimated on the full sample. Panel B: Only weeks with weekends with $PM_{2.5}$ below the city-specific 75th percentile are included. Covariates: Developer and region-by-year-by-quarter fixed effects, developer experience bins, number of public holidays during the workweek, leads of the instrumental variables for the weekend, and flexible weather controls for both the weekend and the workweek (see Equation 4). Standard errors, clustered by city, in parentheses.

and the complexity of PRs developers work on rises by 2.9% of a standard deviation. These findings indicate that developers use low-pollution weekends to compensate for the reduction in work activity and the shift towards easier tasks on high-pollution work days. In other words, developers reallocate work from low to high productivity periods, i.e., from workdays marked by a $PM_{2.5}$ shock towards weekends without air pollution-induced productivity shocks.

To put these effects into perspective, we compare them to the reductions in same-day output due to a $PM_{2.5}$ shock (Table 2). On average, additional work on the weekend makes up for 39% of the reduction in total actions due to a $PM_{2.5}$ shock. On low-pollution weekends, developers even compensate for 58% and 97% of the reduction in commits and total actions, respectively.

To check that the estimates in Table 6 indeed reflect a behavioral response of developers to pollution-induced productivity shocks, we conduct a falsification test. We shift both the weekend and the exposure period forward by two days. The placebo weekend comprises Monday and Tuesday and the placebo exposure period ranges from Wednesday to Sunday of the week before. Since activity levels are low on weekends, productivity shocks during the placebo exposure period, which partially falls on the weekend, should not induce strong compensation during the following week. Moreover, activity is already high on Monday and Tuesday, such that there is not much scope for additional work. Hence, we expect no significant effects of $PM_{2.5}$ exposure. Appendix Table 11 presents the results which confirm this hypothesis. Effects are neither significant in the full sample, nor when considering only placebo weekends with low pollution levels.

In sum, developers work more on weekends to catch up on work not completed due to pollution-induced productivity declines during the workweek. This chance to compensate might allow them to end work early on high-pollution work days. Such a reallocation option could thus also contribute to the absence of effects on work quality in this setting. If developers can end work when their health or cognitive capacity deteriorates and they face an increased risk of committing errors, this will mitigate the impacts of pollution on work quality. At the same time, sacrificing leisure on the weekend, when it is likely most valuable, implies a welfare cost and potentially adverse effects on the work-life balance.

Overall, worker adaptation through changes in task choice and working hours, likely plays an important role in alleviating the effects of $PM_{2.5}$ on output.

6 Heterogeneity and Further Results

Non-Linearity. Exploiting the large variation in air quality in our sample, we investigate whether the effect of pollution on output varies across the range of $PM_{2.5}$ concentrations. To analyze the shape of the dose-response function, we replace $PM_{c,d}$ in equation (1) with a series of dummy variables indicating whether $PM_{2.5}$ concentration falls into a specific bin. Since we do not have enough instruments for running 2SLS, we opt for a more conservative specification with stricter fixed effects for region × date and city × month. These absorb (i) region-wide shocks to developer output on a given date that might be correlated with $PM_{2.5}$ concentration and (ii) seasonal fluctuations in activity and air quality which are allowed to vary across cities. OLS results from this model with either $PM_{2.5}$ in μg/m^3, or the indicator for a $PM_{2.5}$ shock as regressors are presented Columns 4 to 6 of Table 7. Given the finding that the OLS results underestimate the true effects, the results can be interpreted as lower bounds.

Figure 8 displays the estimated impact on total actions when moving from a $PM_{2.5}$ concentration between 7 and 15 μg/m^3 to the respective bin. The baseline bin is chosen such that for each city some observations fall into this range. We find significant reduction in the number of daily actions starting at a level of approximately 70 μg/m^3, but no significant differences for concentrations between the reference bin and 60 μg/m^3. Moving to $PM_{2.5}$ levels below 5 μg/m^3 increases output (point estimate = 0.030, p-value = 0.029). This implies that even in cities with low to moderate levels of $PM_{2.5}$, further improvements in air quality will generate positive effects on worker productivity.

In Appendix Figure 4 we zoom in on the different parts of the function, by grouping the cities into terciles based on average pollution concentration. The most pronounced negative effects of increased pollution on total actions occur in the bottom tercile (mean $PM_{2.5}$ concentration below 8.7 μg/m^3), even at levels below the current regulatory threshold by the U.S. EPA. Given that the OLS estimates likely underestimate the true effects of $PM_{2.5}$ exposure, the results imply relevant economic benefits from complying with the stricter WHO guideline for $PM_{2.5}$. We find no adverse effects of increases in $PM_{2.5}$ in the medium and upper tercile, except for

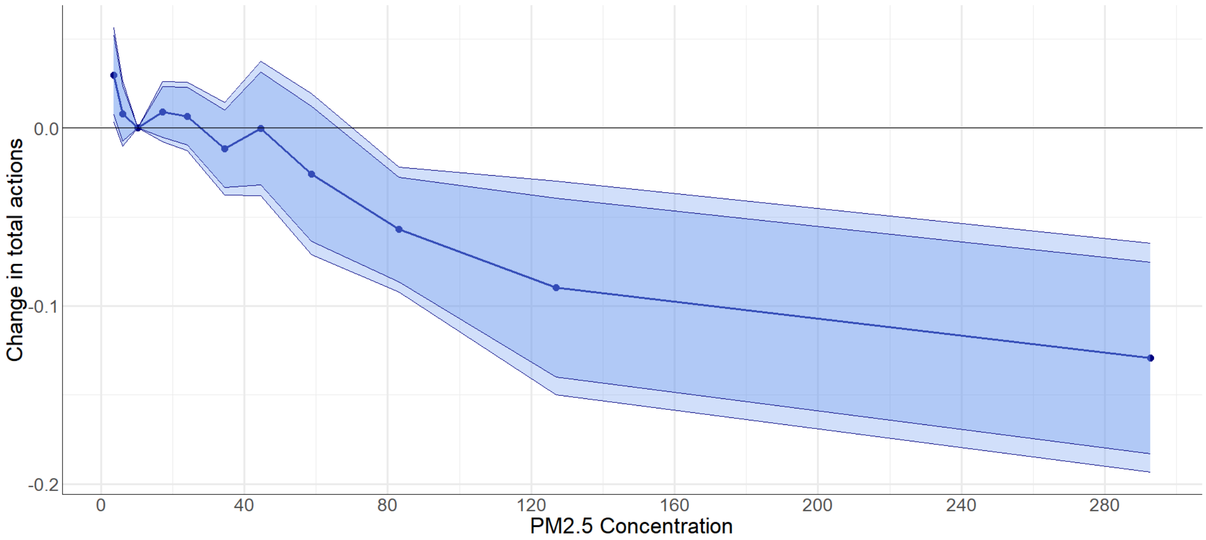


Figure 8: Non-Linear Effects of PM2.5 on Output Quantity

Notes: OLS estimates from regression of total actions on different bins of $PM_{2.5}$ concentrations. Covariates: Weather and holiday controls as in Section 4, region-by-date and city-by-month fixed effects. X-axis: Average $PM_{2.5}$ concentration in each bin in $\mu g/m^3$. Shaded areas: 95% and 90% confidence intervals.

very high concentrations above 40 $\mu g/m^3$. This could be attributed to individuals in cities with higher pollution adopting more avoidance strategies and employing protective measures like air purifiers, which help maintain stable indoor pollution levels except at the most extreme ambient air pollution levels.

Effect Heterogeneity. Next, we analyze heterogeneity in the effect of fine particulate matter by location characteristics to shed light on the distribution of air pollution damages and the potential mechanisms driving the adverse effects on output. We begin by examining the variation in $PM_{2.5}$ effects between locations with different income levels. We group the sample cities based on whether their GDP per capita in 2014 was above or below the *region-specific median.*[13] This approach ensures that our income-based analysis is not confounded by other regional differences correlating with income. The *above median GDP per capita subsample* comprises the cities with relatively high income from each of the seven geographic regions *R*, but includes some cities with lower income than the *below median GDP per capita sample.* Results are reported in Table 7. We find identical absolute effects of a one unit increase in $PM_{2.5}$ concentration on commits and the probability of conducting any action. Relative to the sample mean, effects are slightly larger in the low-GDP sample. For the number of comments and total actions, we find larger effects in low-income cities, in absolute and relative terms. In the high-income sample, effects are not statistically significant. These results align with previous findings that lower-income areas tend to suffer more harm from similar environmental hazards (Colmer et al., 2021; Hsiang et al., 2019).

[13]The main source for local GDP per capita is the OECD metropolitan area database. We assign to each city the value for its respective metro area. Data for cities in non-OECD countries come from national statistical agencies, the OECD regional statistics database, or the World Bank. All values are adjusted for purchasing power parity.

Table 7: Heterogeneity: GDP per capita (IV Estimates)

	Above Median GDP per capita				*Below Median GDP per capita*			
	Actions	*Commits*	*Comments*	*Any Action*	*Actions*	*Commits*	*Comments*	*Any Action*
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$PM_{2.5}$	-0.0018	-0.0014	-0.0002	-0.0002	-0.0036	-0.0014	-0.0012	-0.0002
	(0.0012)	(0.0005)	(0.0006)	(0.0001)	(0.0008)	(0.0005)	(0.0002)	(.00005)
Observations	197,218				200,059			
First stage F-st.	1050				251			
Mean $PM_{2.5}$	14.0				17.9			
Mean dep. var.	2.76	1.29	0.92	0.36	2.48	1.20	0.80	0.35

Notes: Regressor is $PM_{2.5}$ in μg/m^3. Sub-samples based on a city's 2014 GDP per capita compared to the median in its geographic region. Data sources: OECD, World Bank, and national statistical offices. Covariates: See Section 4 or Table 2. Regressions weighted by number of developers in city-month cell. Standard errors, clustered by city, in parentheses.

To assess the potential role of avoidance behavior, we investigate how the effect of $PM_{2.5}$ differs between regions with low and high awareness of air pollution as a significant issue. We use data from the 2020 Pew Research Center International Science Survey to derive a country-wide measure of awareness, namely the share of respondents stating that air pollution is a big problem in their country. 167 of our sample cities in 18 countries are covered by the survey data and we split them into three groups with low, intermediate, and high awareness.[14] Appendix Table 12 presents results from regressions run separately for the three subsamples. In the low awareness group the effect size is twice as large as in the full sample. This may be attributed to lower ownership and utilization of protective devices like air purifiers. Point estimates in the high and intermediate awareness groups are of similar sizes as in the full sample. This pattern suggests that the reduction in output is not due to avoidance behavior, e.g., working from home on high-pollution days, as in such a case we would expect to find monotonically increasing effect magnitude in awareness levels. Given that average pollution concentration does not vary systematically with the level of awareness, nonlinear effects are unlikely to drive the differences.

Lastly, we explore heterogeneity based on the age of the local building stock. Effective exposure to particulate matter is likely lower for individuals in modern buildings with low penetration rates than for those in older, lower-quality buildings given the same outdoor concentration. We use data on the construction period of residential dwellings as a proxy for building stock quality, covering 170 sample cities.[15] We categorize cities into two groups based on the share of dwellings built before 1970, indicating relatively old buildings. We find that the negative effect of $PM_{2.5}$ on total actions is driven by the sample with a higher share of old dwellings, i.e. the cities with likely higher effective exposure (Panel B of Appendix Table 12).

[14]US cities form the intermediate awareness sample, with 63% deeming air pollution a big problem. Countries where a larger (smaller) share of respondents holds this view, are assigned to the high (low) awareness sample.

[15]The data are collected from the American Community Survey for metropolitan areas, the EU Building Stock Observatory (country-level), the Federal Statistical Office of Switzerland (canton-level), the Statistics Bureau of Japan (prefecture-level), Statistics Canada (province-level), and Statistics Norway (municipality-level).

Table 8: Analysis at the Monthly Level (IV Estimates)

	Actions		*Commits*		*Comments*		*Follower Growth*	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$PM_{2.5}$	-0.131		-0.059		-0.040		-0.00004	
	(0.027)		(0.021)		(0.008)		(0.00002)	
$PM_{2.5}$ shock		-3.980		-1.812		-1.673		-0.0022
		(1.961)		(0.9756)		(0.9226)		(0.0007)
First stage F-stat.	63	26	63	26	63	26	65	26
Observations	529,682		529,682		529,682		511,361	
Mean dep. var.	81.4		38.5		26.9		0.0073	

Notes: The regressor is average monthly $PM_{2.5}$ in µg/m^3 (odd-numbered columns), or a dummy indicating that monthly $PM_{2.5}$ exceeds the city-specific mean by at least one city-specific standard deviation (even-numbered columns). Covariates: Developer and region-by-year-by-month fixed effects, bins for developer's experience on GitHub, third-order polynomials in average monthly temperature, precipitation, relative humidity, and wind speed, number of holidays and days with heavy wildfire smoke. Effects of weather controls and holidays can vary across regions. Regressions in Columns (7)-(8) also control for a second-order polynomial of the number of followers at the start of the month. Standard errors, clustered by city, in parentheses.

This suggests that the main results are attributable to physiological responses to air pollution, rather than behavioral changes or avoidance behavior.

Monthly Level. Next, we quantify the effect of monthly $PM_{2.5}$ on monthly output and on the growth rate of the number of the developer's followers as a summary measure for the quantity, quality, and relevance of their work on GitHub. With this, we aim to analyze pollution effects net of short-run compensation responses and to investigate whether we find evidence for more *long-run* consequences of pollution exposure proxied by follower growth. We run 2SLS regressions at the developer×month level, again using two distinct regressors, average monthly $PM_{2.5}$ concentration in µg/m^3, or an indicator that takes value one if this concentration exceeds the city-specific mean by at least one city-specific standard deviation. We instrument these variables with average monthly inversion strength, interacted with indicators for the 25 city-groups g.

Table 8 shows that a 1 µg/m^3 increase in monthly $PM_{2.5}$ reduces monthly actions by 0.13, 0.16% relative to the mean. The implied reduction in daily actions is 0.0043, and thus slightly larger than the effect found in the analysis at the daily level (0.0030). Again, this effect is driven by reductions in both commits and comments. $PM_{2.5}$ also reduces follower growth by 0.5% relative to the mean. In line with our results at the daily level, we see that months with unusually high pollution levels largely drive the effects, reducing the output measures by 4.9% to 6.2%, and the growth rate of the number of followers by 30%, relative to months with better air quality. In sum, air pollution negatively affects developers' output also over more aggregated periods. It even hinders the growth of reputation in the tech community, potentially impacting developers' career prospects.

Robustness. Table 13 shows that our main results for work quantity, work time, and task choice are robust to different specifications of the first stage. These changes include replacing inversion strength, $\Delta T_{c,d}$, by an inversion indicator, $\mathbf{1}\{\Delta T_{c,d} > 0\}$, as well as varying the number of city-groups g across which the effects of inversion strength can differ in the first stage to 15 or 50 (baseline specifications uses 25 groups).

In Table 14, we test alternative functional forms in the second stage. When we apply the inverse hyperbolic sine transformation to our measures of work output, the direction and statistical significance of the baseline results persist, but this specification implies somewhat smaller effect magnitudes. Using $PM_{2.5}$ in logs as regressor yields the same pattern for second-stage effects on work quantity, time of last action, and task complexity as the baseline model. We can also replicate our results with an alternative, policy-based indicator for unusually high-pollution days which takes value one on when $PM_{2.5}$ exceeds the relevant air quality standard in the respective country.

Lastly, we demonstrate that the results are overall robust to varying the included sets of fixed effects, absorbing common shocks at different geographic and temporal levels, as well as varying the included weather controls (Appendix Figures 5 to 8).

7 Conclusion

How do environmental conditions affect workers in jobs that form the backbone of the modern knowledge economy? In our paper, we provide insights on this question, which is highly relevant as such jobs are expected to become even more widespread as digitalization and automation continue to change the world of work.

Using rich data on GitHub activities, we show that air pollution reduces daily output in a global sample of software developers. This decline is mostly driven by reductions in individual coding activity and work on new tasks, while collaborative activity in response to others' work is less affected. Our estimates are at the lower end of air pollution effects found in less flexible and less collaborative occupations. Moreover, we find only minor deteriorations in output quality. Due to the high value generated by software developers, the implied monetary loss is nevertheless economically relevant and comparable to findings for manual occupations. Developers exploit the flexibility of their work setting to adapt to environmental shocks. When $PM_{2.5}$ increases, they shift to less complex tasks and reallocate work from high-pollution, low-productivity workdays to low-pollution, high-productivity weekends. These adjustments likely help to alleviate effects on output quantity and quality, but they also imply additional welfare cost of air pollution not captured by changes in output due to forgone leisure time on weekends and negative impacts on work-life balance.

While we use data on software developers using GitHub as part of their formal work, we believe that the findings generalize to other occupations that offer flexible schedules and discretion in task choice and require similar skills like problem-solving, attention to detail, programming, and teamwork. This applies to many knowledge workers, including business analysts or researchers. Furthermore, since our sample covers 47 countries, the estimates in this study are not specific to a certain firm or country context.

Based on this, we derive estimates of the monetary benefits from reducing $PM_{2.5}$ concentration permanently by one $\mu g/m^3$ in terms of productivity gains among knowledge workers. Extrapolating to all US workers in the occupation group 'Computer and Mathematical Workers' and all ICT professionals in the European Union suggests annual benefits of $360m (US) and $600m (EU), respectively.[16] Hence, our findings have important policy implications. When setting air pollution standards, regulators may want to factor in the growing evidence on the economic benefits of pollution reductions through productivity gains. Importantly, we find adverse effects of $PM_{2.5}$ on output even below the current regulatory standards in the European Union and the US. While we find slightly smaller marginal effects in highly polluted locations, the substantially greater concentrations of $PM_{2.5}$ in developing countries like India and Bangladesh, compared to the US, could pose a challenge for growing their software industries.

Our findings on how software developers adjust work patterns also have interesting implications for the organization of work within firms: Highlighting the difficulty of certain tasks and granting flexibility in working hours might help workers to better adapt to idiosyncratic productivity shocks and mitigate the total impact on team or firm performance.

References

Adhvaryu, A., N. Kala, and A. Nyshadham (2022): "Management and Shocks to Worker Productivity," *Journal of Political Economy*, 130, 1–47.

Angelici, M. and P. Profeta (2020): "Smart-working: Work flexibility without constraints," Working Paper, CESifo.

Arceo-Gomez, E. O., R. Hanna, and P. Oliva (2016): "Does the Effect of Pollution on Infant Mortality Differ Between Developing and Developed Countries? Evidence from Mexico City," *Economic Journal*, 126, 257–280.

Archsmith, J., A. Heyes, and S. Saberian (2018): "Air Quality and Error Quantity: Pollution and Performance in a High-Skilled, Quality-Focused Occupation," *Journal of the Association of Environmental and Resource Economists*, 5, 827–863.

Autor, D. H., F. Levy, and R. J. Murnane (2003): "The skill content of recent technological change: An empirical exploration," *The Quarterly Journal of Economics*, 118, 1279–1333.

Autor, D. H. and B. Price (2013): "The Changing Task Composition of the US Labor Market: An Update of Autor, Levy, and Murnane (2003)," Working Paper.

[16] Combining our results with the estimated monetary value of commits and PRs implies that a one unit decrease in $PM_{2.5}$ increases daily output value by $0.213. With 4,654,750 'Computer and Mathematical Workers' in the US in 2021 (Bureau of Labor Statistics, 2022), the annual gain in output value of a permanent one $\mu g/m^3$ pollution reduction for this group is thus $\$0.213 \times 4{,}654{,}750 \times 365$ days = $360m. We compute the value for the EU analogously, based on 7,843,000 ICT professionals in 2020 (Cedefop, 2022).

Babadjouni, R. M., D. M. Hodis, R. Radwanski, R. Durazo, A. Patel, Q. Liu, and W. J. Mack (2017): "Clinical effects of air pollution on the central nervous system; a review," *Journal of Clinical Neuroscience*, 43, 16–24.

Bassi, V., M. E. Kahn, N. Lozano Gracia, T. Porzio, and J. Sorin (2021): "Pollution in Ugandan Cities: Do Managers Avoid it or Adapt in Place?" Working Paper.

Baylis, P. (2020): "Temperature and temperament: Evidence from Twitter," *Journal of Public Economics*, 184, 104161.

Beckmann, M., T. Cornelissen, and M. Kräkel (2017): "Self-managed working time and employee effort: Theory and evidence," *Journal of Economic Behavior & Organization*, 133, 285–302.

Borgschulte, M., D. Molitor, and E. Zou (forthcoming): "Air Pollution and the Labor Market: Evidence from Wildfire Smoke," *The Review of Economics and Statistics*.

Bresnahan, T. F., E. Brynjolfsson, and L. M. Hitt (2002): "Information technology, workplace organization, and the demand for skilled labor: Firm-level evidence," *The Quarterly Journal of Economics*, 117, 339–376.

Bureau of Labor Statistics (2021): "Software Developers, Quality Assurance Analysts, and Testers," in *Occupational Outlook Handbook*, U.S. Department of Labor.

——— (2022): "Occupational Employment and Wage Statistics," https://www.bls.gov/oes/current/oes_nat.htm, accessed: 2022-11-07.

Burke, M., S. Heft-Neal, J. Li, A. Driscoll, P. W. Baylis, M. Stigler, J. Weill, J. Burney, J. Wen, M. Childs, and C. Gould (2022): "Exposures and Behavioral Responses to Wildfire Smoke," *Nature Human Behavior*.

Cedefop (2022): "Employed population by occupation and sector," https://www.cedefop.europa.eu/en/tools/skills-intelligence/employed-population-occupation-and-sector?year=2020&country=EU&occupation=#1, accessed: 2022-11-07.

Chang, T., J. Graff Zivin, T. Gross, and M. Neidell (2016): "Particulate Pollution and the Productivity of Pear Packers," *American Economic Journal: Economic Policy*, 8, 141–169.

Chang, T. Y., J. Graff Zivin, T. Gross, and M. Neidell (2019): "The Effect of Pollution on Worker Productivity: Evidence from Call Center Workers in China," *American Economic Journal: Applied Economics*, 11, 151–172.

Colmer, J., D. Lin, S. Liu, and J. Shimshack (2021): "Why are pollution damages lower in developed countries? Insights from high-Income, high-particulate matter Hong Kong," *Journal of Health Economics*, 79, 102511.

Currie, J., L. Davis, M. Greenstone, and R. Walker (2015): "Environmental Health Risks and Housing Values: Evidence from 1,600 Toxic Plant Openings and Closings," *American Economic Review*, 105, 678–709.

Currie, J. and M. Neidell (2005): "Air Pollution and Infant Health: What Can We Learn from California's Recent Experience?" *The Quarterly Journal of Economics*, 120, 1003–1030.

Deng, G., Z. Li, Z. Wang, J. Gao, Z. Xu, J. Li, and Z. Wang (2017): "Indoor/outdoor relationship of PM2.5 concentration in typical buildings with and without air cleaning in Beijing," *Indoor and Built Environment*, 26, 60–68.

Donald, S. G. and K. Lang (2007): "Inference with Difference-in-Differences and Other Panel Data," *Review of Economics and Statistics*, 89, 221–233.

Ebenstein, A., V. Lavy, and S. Roth (2016): "The Long-Run Economic Consequences of High-Stakes Examinations: Evidence from Transitory Variation in Pollution," *American Economic Journal: Applied Economics*, 8, 36–65.

Fu, S., V. B. Viard, and P. Zhang (2021): "Air Pollution and Manufacturing Firm Productivity: Nationwide Estimates for China," *The Economic Journal*, 131, 3241–3273.

Graff Zivin, J. and M. Neidell (2012): "The Impact of Pollution on Worker Productivity," *American Economic Review*, 102, 3652–3673.

——— (2014): "Temperature and the Allocation of Time: Implications for Climate Change," *Journal of Labor Economics*, 32, 1–26.

He, J., H. Liu, and A. Salvo (2019): "Severe Air Pollution and Labor Productivity: Evidence from Industrial Towns in China," *American Economic Journal: Applied Economics*.

Heyes, A., N. Rivers, and B. Schaufele (2019): "Pollution and Politician Productivity: The Effect of PM on MPs," *Land Economics*, 95, 157–173.

Hoffmann, B. and J. P. Rud (2022): "Exposure or Income? The Unequal Effects of Pollution on Daily Labor Supply," Working Paper, Inter-American Development Bank.

Hsiang, S., P. Oliva, and R. Walker (2019): "The Distribution of Environmental Damages," *Review of Environmental Economics and Policy*, 13, 83–103.

Huang, J., N. Xu, and H. Yu (2020): "Pollution and Performance: Do Investors Make Worse Trades on Hazy Days?" *Management Science*, 66, 4455–4476.

Isen, A., M. Rossin-Slater, and W. R. Walker (2017): "Every Breath You Take—Every Dollar You'll Make: The Long-Term Consequences of the Clean Air Act of 1970," *Journal of Political Economy*, 125, 848–902.

Jans, J., P. Johansson, and J. P. Nilsson (2018): "Economic status, air quality, and child health: Evidence from inversion episodes," *Journal of Health Economics*, 61, 220–232.

Kahn, M. E. and P. Li (2020): "Air pollution lowers high skill public sector worker productivity in China," *Environmental Research Letters*, 15, 084003.

Karagulian, F., C. A. Belis, C. F. C. Dora, A. M. Prüss-Ustün, S. Bonjour, H. Adair-Rohani, and M. Amann (2015): "Contributions to cities' ambient particulate matter (PM): A systematic review of local source contributions at global level," *Atmospheric Environment*, 120, 475–483.

Krebs, B. and S. Luechinger (forthcoming): "Air Pollution, Cognitive Performance, and the Role of Task Proficiency," *Journal of the Association of Environmental and Resource Economists*.

Künn, S., J. Palacios, and N. Pestel (2023): "Indoor Air Quality and Strategic Decision-Making," *Management Science*, 69, 5354–5377.

La Nauze, A. and E. R. Severnini (2021): "Air Pollution and Adult Cognition: Evidence from Brain Training," Working Paper 28785, National Bureau of Economic Research.

Lederer, A. M., P. M. Fredriksen, B. N. Nkeh-Chungag, F. Everson, H. Strijdom, P. De Boever, and N. Goswami (2021): "Cardiovascular effects of air pollution: current evidence from animal and human studies," *American Journal of Physiology-Heart and Circulatory Physiology*, 320, H1417–H1439, pMID: 33513082.

LoPalo, M. (2023): "Temperature, Worker Productivity, and Adaptation: Evidence from Survey Data Production," *American Economic Journal: Applied Economics*, 15, 192–229.

Mas, A. and A. Pallais (2020): "Alternative Work Arrangements," *Annual Review of Economics*, 12, 631–658.

McDermott, G. and B. Hansen (2021): "Labor Reallocation and Remote Work During COVID-19: Real-time Evidence from GitHub," Working Paper 29598, National Bureau of Economic Research.

Menon, S., A. Salvatori, and W. Zwysen (2020): "The effect of computer use on work discretion and work intensity: Evidence from Europe," *British Journal of Industrial Relations*, 58, 1004–1038.

Mullins, J. T. and C. White (2019): "Temperature and mental health: Evidence from the spectrum of mental health outcomes," *Journal of Health Economics*, 68, 102240.

Muro, M., S. Liu, J. Whiton, and S. Kulkarni (2017): "Digitalization and the American workforce," Working Paper, Brookings India.

Neidell, M., J. G. Zivin, M. Sheahan, J. Willwerth, C. Fant, M. Sarofim, and J. Martinich (2021): "Temperature and Work: Time Allocated to Work under Varying Climate and Labor Market Conditions," *PLOS ONE*, 16, e0254224.

Park, R. J. (2020): "Hot temperature and high stakes performance," *Journal of Human Resources*.

Pope, C. A. (2000): "Epidemiology of Fine Particulate Air Pollution and Human Health: Biologic Mechanisms and Who's at Risk?" *Environmental Health Perspectives*, 108, 713–723.

Sager, L. (2019): "Estimating the effect of air pollution on road safety using atmospheric temperature inversions," *Journal of Environmental Economics and Management*, 98, 102250.

Sarmiento, L. (2022): "Air pollution and the productivity of high-skill labor: evidence from court hearings," *The Scandinavian Journal of Economics*, 124, 301–332.

Somanathan, E., R. Somanathan, A. Sudarshan, and M. Tewari (2021): "The impact of temperature on productivity and labor supply: Evidence from Indian manufacturing," *Journal of Political Economy*, 129, 1797–1827.

Tan, X., M. Zhou, and Z. Sun (2020): "A first look at good first issues on github," in *Proceedings of the 28th ACM Joint Meeting on European Software Engineering Conference and Symposium on the Foundations of Software Engineering*, 398–409.

Appendix (For Online Publication)

A Additional Tables

Table 1: Characteristics of High-Skill Occupations and Software Development

	Freedom to Make Decisions		Structured versus Unstructured Work		Work With Work Group or Team	
	All high-skill occupations	Software developers	All high-skill occupations	Software develop.	All high-skill occupations	Software develop.
1	0.4	0	0.6	0	1.79	0
2	2.6	3.1	2.3	2.4	4.4	5.9
3	10.5	29.1	11.3	28.1	11.0	2.7
4	35.6	38.2	39.8	45.0	30.5	9.2
5	50.9	29.6	46.0	24.6	52.4	82.3

Notes: Based on data from O*NET Database Version 25.0. Work Contexts Table. All high-skill occupations refers to occupations in Job Zones 4 and 5. Software developers refers to occupation 15-1132.00 ("Software Developers, Applications"). Categories: 1 = not important at all/no freedom; 2 = Fairly important/very little freedom; 3 = Important/Limited freedom; 4= Very Important/Some freedom; 5 = Extremely important / A lot of freedom

Table 2: Labels Indicating *Easy* Issues

good first issues	good first bug	good-first
documentation	polish	cleanup
simple	easy	small
trivial	minor	**help wanted**
junior job	newcomer	starter
beginner	newbie	novice
low hanging	low-hanging	

Notes: If a label contains any of these terms, the issue is classified as "easy". Bolt text indicates GitHub default labels.

Table 3: Description of Outcome Variables

Domain	Concept	Variable	Details
Output Quantity	Total output quantity	Actions	Sum of number of commits, comments on issues, PRs and commits, PRs opened, PRs closed, issues opened, closed and reopened
	Coding activity	Commits	Number of commits
	Interactive activity	Comments	Sum of number of comments written on issues, PRs and commits
Output Quality	Proposed code changes that get accepted	PR Acceptance Rate	$\frac{\text{PRs opened that got merged}}{\text{all PRs opened}}$
	Deficient commits	Share commits reverted	$\frac{\text{Commits that got reverted}}{\text{all commits}}$
Task choice	Average PR complexity	Lines added per PR	Average number of lines of code added in PRs opened or closed
		Files changed per PR	Average number of code files changed in PRs opened or closed
		Lines deleted per PR	Average number of lines of code deleted in PRs opened or closed
	Easy tasks among issue events	Share easy issue events	(#easy issues opened + #easy issues closed + #comments written on easy issues)/ (#issues opened + #issues closed + #comments written on issues)
Working hours	Evening activity	Time last action	Minute of final action of the day
	Weekend Activity	Actions	Number of total actions conducted on the weekend

Notes: The Table displays information on the outcome variables we use, how they are constructed, and what they measure.

Table 4: Sources of Air Quality Data

Geographic Area	**Data Source**
United States	U.S. Environmental Protection Agency (EPA)
Canada	Canadian National Air Pollution Surveillance (NAPS) Program
Mexico City	Gobierno de la Ciudad de México
Europe	European Environment Agency (EEA)
Russia, Ukraine, Belarus, Turkey, Israel	Copernicus Atmosphere Monitoring Service (CAMS)
China	National Environmental Monitoring Centre
Chennai, Mumbai, New Delhi, Dhaka, Hyderabad	US Embassies (AirNow.gov)
Bengaluru	Central Pollution Control Board (CPCB)
Japan	National Institute for Environmental Studies
Hong Kong	Hong Kong Environmental Protection Department
Singapore	National Environment Agency
South Korea	Air Korea
Taiwan	Environmental Protection Administration
Australia	New South Wales Department of Planning and Environment Victorian Government open data portal Queensland Government open data portal South Australian Government Data Directory
New Zealand	Stats NZ Tatauranga Aotearoa

Notes: Data sources for data on $PM_{2.5}$. Airbase, the EEA's database on air pollution, contains monitor data for 33 countries, including all EU members, as well as further EEA member and cooperating countries, e.g., Switzerland, Norway and Serbia.

Table 5: Developer-by-date observations and inversion frequency by geographic regions R

Region R	Observations	Share of Total Observations (%)	Inversion Frequency (%)
Oceania	502,942	3.1	35
Northern America	7,374,420	45.6	50
Northern Europe	1,809,795	11.2	35
Western Europe	2,297,385	14.2	53
Southern Europe	474,070	2.9	51
Eastern Europe	1,024,998	6.3	57
Asia	2,691,543	16.6	30

Notes: The table shows the distribution of observations in the developer × date panel described in section 3.2 across geographic regions R on the left. On the right, it shows the share of all city × date observations with a nighttime inversion by geographic region R.

Table 6: Effect of $PM_{2.5}$ on Quantity of Issue and Pull Request Actions

	PRs closed (1)	*PRs opened* (2)	*Issues closed* (3)	*Issues opened* (4)
Panel A.				
$PM_{2.5}$	-0.00013	-0.00016	-0.00009	-0.00023
	(0.00005)	(0.00004)	(0.00004)	(0.00004)
First stage F-stat.	160	160	160	160
% change	-0.08	-0.11	-0.08	-0.23
Panel B.				
$PM_{2.5}$ shock	-0.0037	-0.0066	0.0051	-0.0077
	(0.0064)	(0.0058)	(0.0042)	(0.0039)
First stage F-stat.	76	76	76	76
% change	-2.3	-4.4	4.6	-7.7
Mean dep. var.	0.16	0.15	0.12	0.10
Observations	397,277	397,277	397,277	397,277

Notes: 2SLS estimates of the parameter β in Equation (1). The regressor of interest is $PM_{2.5}$ concentration in $\mu g/m^3$ (Panel A) or a binary $PM_{2.5}$ shock variable as in Equation (3) (Panel B). Covariates include fixed effects and flexible weather controls as described in Table 2. Regressions are weighted by the number of active workers in a city-month cell. Standard errors clustered at the city level are reported in parentheses.

Table 7: OLS Results for Work Quantity

	Actions (1)	*Commits* (2)	*Comments* (3)	*Actions* (4)	*Commits* (5)	*Comments* (6)
Panel A.						
$PM_{2.5}$	−0.00047	−0.00019	−0.00017	−0.00034	−0.00014	−0.00014
	(0.00013)	(0.00008)	(0.00007)	(0.00016)	(0.00011)	(0.00004)
Panel B.						
$PM_{2.5}$ shock	−0.0068	−0.0060	−0.0004	−0.0083	−0.0063	−0.0013
	(0.0105)	(0.0051)	(0.0038)	(0.0109)	(0.0053)	(0.0041)
Observations	397,277	397,277	397,277	397,277	397,277	397,277
City FE	✓	✓	✓			
Region×Day-of-Week FE	✓	✓	✓			
Region×Year-Month FE	✓	✓	✓			
Region×Date FE				✓	✓	✓
City×Month FE				✓	✓	✓

Notes: OLS estimates of the parameter β in Equation (1). The regressor of interest is $PM_{2.5}$ concentration in $\mu g/m^3$ (Panel A) or a binary $PM_{2.5}$ shock variable as in Equation (3) (Panel B). Covariates: flexible weather controls as described in Section 4. Included fixed effects are displayed in the bottom part of the table. Regressions are weighted by the number of active workers in a city-month cell. Standard errors clustered at the city level are reported in parentheses.

Table 8: Lagged Effects of Inversion Strength on Output Quantity

	Actions (1)	*Commits* (2)	*Comments* (3)	*Any Actions* (4)
$Invstrength_{c,d}$	-0.0040	-0.0018	-0.0013	-0.0002
	(0.0010)	(0.0005)	(0.0006)	(0.0001)
$Invstrength_{c,d-1}$	0.0003	0.0001	-0.0001	0.00004
	(0.0013)	(0.0006)	(0.0006)	(0.0001)
$Invstrength_{c,d-2}$	0.0003	0.00002	0.0002	0.0001
	(0.0010)	(0.0004)	(0.0005)	(0.0001)
$Invstrength_{c,d-3}$	0.0018	0.0007	0.0005	0.00004
	(0.0011)	(0.0006)	(0.0004)	(0.0001)
Observations	417,116	417,116	417,116	417,116
'First stage effect' of $Invstrength_{c,d}$ on $PM_{2.5}$			0.869 (0.114)	

Notes: OLS estimates of the outcomes displayed at the top on inversion strength in degrees Celsius on the same day, as well as three lags. Covariates: fixed effects as described in Table 2 and third order-polynomials in mean daily temperature, precipitation, wind speed and relative humidity as well as three lags of the weather controls. Effects of the weather controls can vary across world regions *R*. At the bottom of the table, the coefficient from a regression of $PM_{2.5}$ in µg/m^3 on inversion strength (with fixed effects and weather controls) is presented. Regressions are weighted by the number of active workers in a city-month cell. Standard errors clustered at the city level are reported in parentheses.

Table 9: Effect of $PM_{2.5}$ on PRs opened and closed with GHArchive and GHTorrent data

	PRs opened (GHA) (1)	*PRs closed (GHA)* (2)	*PRs opened (GHT)* (3)	*PRs closed (GHT)* (4)
$PM_{2.5}$	-0.0002	-0.0001	-0.0002	-0.0001
	(0.00003)	(0.00005)	(0.00004)	(0.00005)
	[<0.00001]	[0.008]	[<0.00001]	[0.018]
First stage F-stat.	165	165	165	165
Observations	337,008	337,008	337,008	337,008

Notes: 2SLS estimates of the parameter β in equation (1), where the outcome is the number of pull requests (PRs) opened or closed, respectively. This is measured in GHArchive data in columns (1) to (2) and GHTorrent data in column (3) to (4). The regressor of interest is $PM_{2.5}$ concentration in $\mu g/m^3$. Covariates are as described in Table 2. The sample period is 2015 to May 2019. Regressions are weighted by the number of active workers in a city- month cell. Standard errors clustered at the city level are reported in parentheses. P-values are reported in brackets.

Table 10: Pull Request Complexity Index

	PR Complexity	
	(1)	(2)
$PM_{2.5}$ (in µg/m^3)	-0.0003	
	(0.0002)	
$PM_{2.5}$ shock		-0.0273
		(0.0158)
Observations	1,750,825	1,750,825
First stage F-stat.	215	68

Notes: 2SLS estimates of the effect of $PM_{2.5}$ on the complexity of PRs a developer opened or closed on the current day, estimated at the developer × day level. The dependent variable is a PR complexity index that is computed as the average of the mean number of new lines of code added, number of lines of code deleted and number of code files changed per PR, after standardizing each variable. The complexity index is divided by its standard deviation. Covariates: flexible weather controls as described in Table 2, as well as developer, region-by-day-of-week, and region-by-year-by-month fixed effects. Regressions are based on GHArchive data. The sample period is 2015 to May 2019. Standard errors clustered at the city level are reported in parentheses.

Table 11: Placebo Test: Effect of $PM_{2.5}$ Wednesday to Sunday on Work Activity Monday to Tuesday

	Actions	*Commits*	*Comments*	*Any action*
	(1)	(2)	(3)	(4)
Panel A: Full Sample				
Days with $PM_{2.5}$ shocks	0.0237	0.0033	0.0010	0.0003
	(0.0499)	(0.0244)	(0.0207)	(0.0018)
Observations	2,241,316	2,241,316	2,241,316	2,241,316
Panel B: low PM Weekends only				
Days with $PM_{2.5}$ shocks	-0.0311	-0.0041	-0.0373	-0.0035
	(0.0682)	(0.0343)	(0.0296)	(0.0029)
Observations	1,505,973	1,505,973	1,505,973	1,505,973

Notes: 2SLS estimates of the parameter β in a placebo version of equation 4. Outcomes are the sum of all actions, commits and comments made between Monday and Tuesday, the placebo weekend. The regressor of interest is the number of days characterized by a $PM_{2.5}$ shock during the placebo workweek, i.e. the previous five days. Regressions control for developer and region-by-year-by-quarter fixed effects, developer experience bins, the number of public holidays during the workweek, and the leads of the instrumental variables for the placebo weekend. Further covariates are flexible weather controls for the placebo exposure period and the placebo weekend. Standard errors clustered at the city level are reported in parentheses.

Table 12: Heterogeneity: Awareness and Building Stock Age

	Dependent Variable: Actions			
Panel A. Awareness				
	All	*High*	*Intermediate*	*Low*
	(1)	(2)	(3)	(4)
$PM_{2.5}$	-0.0033	-0.0029	-0.0033	-0.0080
	(0.0007)	(0.0006)	(0.0059)	(0.0019)
Observations	307,313	84,079	141,036	82,198
First stage F-stat.	259	357	161	555
Share *Air Pollution is Big Problem*	65.0%	79.7%	63.1%	52.8%
Mean $PM_{2.5}$	12.0	19.3	8.3	11.0
Mean dep. var.	2.8	2.5	3.0	2.7
Panel B. Building Stock Age				
	All	*Above Median*	*Below Median*	
	(1)	(2)	(3)	
$PM_{2.5}$	-0.0046	-0.0079	-0.0023	
	(0.0026)	(0.0023)	(0.0044)	
Observations	309,966	150,236	159,730	
First stage F-stat.	161	184	265	
Share modern buildings	27%	17%	36%	
Share old buildings	46%	59%	33%	
Mean $PM_{2.5}$	10.3	10.7	9.7	
Mean dep. var.	2.93	2.87	3.00	

Notes: In Panel A, the sample used in Column (1) includes all 167 cities covered by the Pew Research Center International Science Survey. Results in Columns (2) to (4) are estimated on subsamples formed based on country-level awareness of air pollution, measured by the share of respondents stating that air pollution is a big problem in the Pew Survey. In Panel B, the sample used in Column (1) includes all 170 cities covered by data on building stock age. Results in Columns (2) to (3) are estimated on subsamples formed based on the share of dwellings built before 1970, which are defined as old buildings. Modern buildings are those built after 1990. Regressions weighted by number of developers in city-month cell. Standard errors, clustered by city, in parentheses.

Table 13: Robustness: First Stage Specification

	Actions (1)	*Commits* (2)	*Comments* (3)	*Any Action* (4)	*Time of Last Action* (5)	*Lines added per PR* (6)	*Files changed per PR* (7)
Panel A.							
$PM_{2.5}$	-0.0044	-0.0023	-0.0013	-0.0003	-0.1490	-0.0028	-0.0014
	(0.0006)	(0.0004)	(0.0002)	(0.00006)	(0.036)	(0.0013)	(0.0004)
First stage F-stat.	83	83	83	83	84	146	146
Panel B.							
$PM_{2.5}$ shock	-0.1358	-0.1018	-0.0365	-0.0091	-3.199	-0.0242	-0.0324
	(0.0744)	(0.0353)	(0.0328)	(0.0056)	(3.503)	(0.0281)	(0.0160)
First stage F-stat.	49	49	49	49	50	59	59
IV-Specification	Inversion Indicator $\times$ 25 city groups						
Observations	397,277	397,277	397,277	397,277	339,872	183,749	183,749
Panel C.							
$PM_{2.5}$	-0.0029	-0.0015	-0.0008	-0.0002	-0.0886	-0.0021	-0.0011
	(0.0006)	(0.0004)	(0.0002)	(0.00004)	(0.0374)	(0.0008)	(0.0003)
First stage F-stat.	750	750	750	750	701	1040	1040
Panel D.							
$PM_{2.5}$ shock	-0.1026	-0.0682	-0.0247	-0.0096	-3.998	-0.0337	-0.0284
	(0.0688)	(0.0300)	(0.0303)	(0.0048)	(2.491)	(0.0239)	(0.0118)
First stage F-stat.	198	198	198	198	144	167	167
IV-Specification	Inversion Strength $\times$ 50 city groups						
Observations	397,277	397,277	397,277	397,277	339,872	183,749	183,749
Panel E.							
$PM_{2.5}$	-0.0031	-0.0016	-0.0009	-0.0002	-0.0896	-0.0022	-0.0011
	(0.0006)	(0.0004)	(0.0002)	(4.3×10^{-5})	(0.0379)	(0.0007)	(0.0003)
First stage F-stat.	153	153	153	153	203	373	373
Panel F.							
$PM_{2.5}$ shock	-0.1298	-0.0761	-0.0393	-0.0115	-5.110	-0.0410	-0.0349
	(0.0658)	(0.0305)	(0.0291)	(0.0048)	(2.547)	(0.0264)	(0.0144)
First stage F-stat.	86	86	86	86	79	80	80
IV-Specification	Inversion Strength $\times$ 15 city groups						
Observations	397,277	397,277	397,277	397,277	339,872	183,749	183,749

Notes: 2SLS estimates of the parameter β in Equation (1). In Panels A, C and E, the regressor of interest is $PM_{2.5}$ concentration in $\mu g/m^3$. In Panel B, D and F, an indicator for a $PM_{2.5}$ shock is used instead. Relative to specifications underlying results in Table 2, the first stage model is changed: In Panels A and B, instruments are the interactions of a dummy variable for a thermal inversion occurring with indicators for 25 distinct city groups g. In Panels C to F, the first stage specification is as in Equation (2), but we form 50 city-groups g (C and D) or 15 city-groups g (E and F) instead of 25 groups (baseline). Covariates as described in Table 2. Regressions are weighted by the number of active workers in a city during the current month. Standard errors clustered at the city level are reported in parentheses.

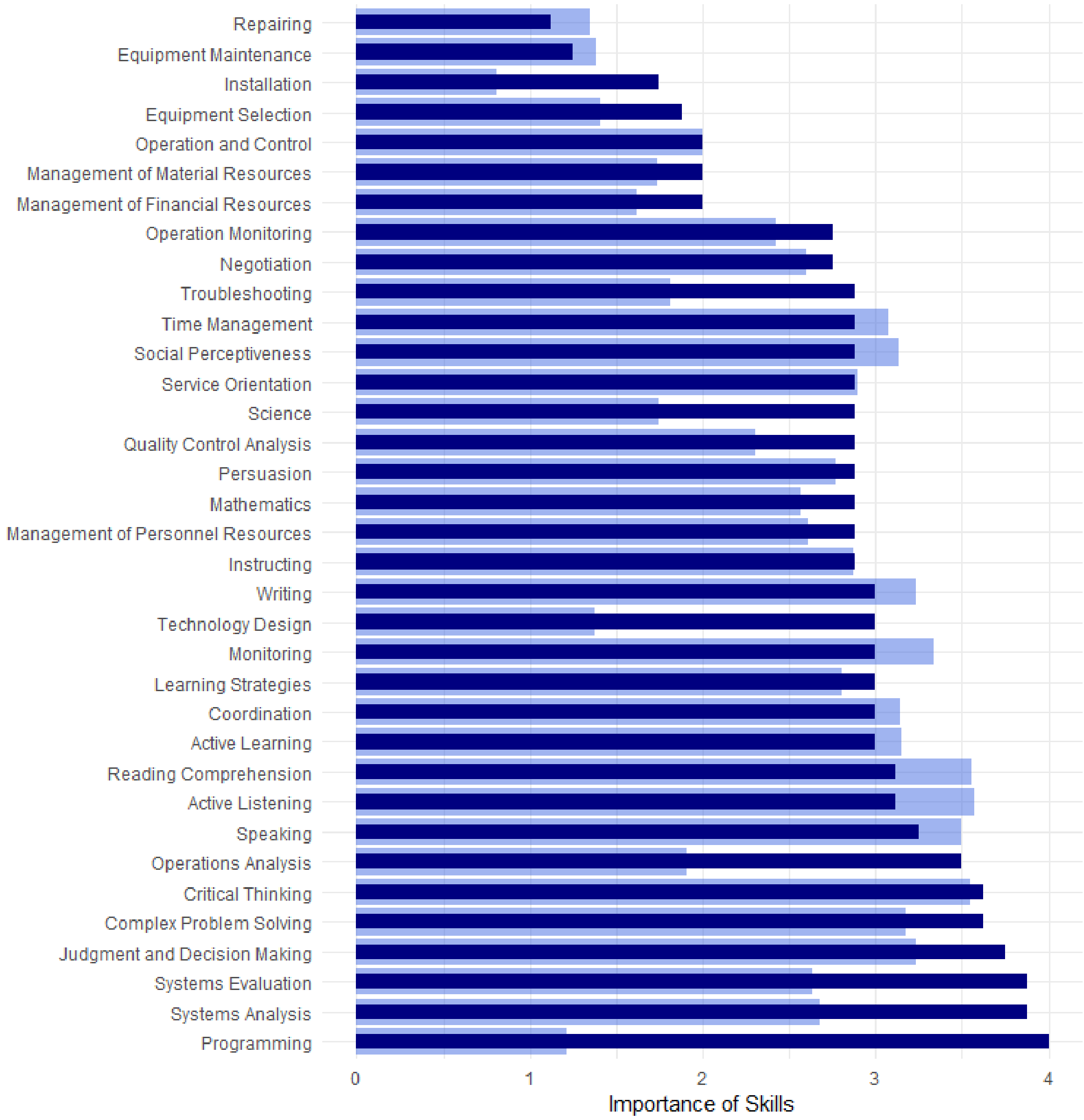
Table 14: Robustness: Second Stage Specification

	Actions (1)	*Commits* (2)	*Comments* (3)	*Any Action* (4)	*Time of Last Action* (5)	*Lines added per PR* (6)	*Files changed per PR* (7)
Panel A. Inv. Hyperbolic Sine Transformation							
$PM_{2.5}$	−0.0006	−0.0005	−0.0003				
	(0.0001)	(0.0001)	(0.0001)				
First stage F-stat.	160	160	160				
Panel B. Inv. Hyperbolic Sine Transformation							
$PM_{2.5}$ shock	−0.0250	−0.0185	−0.0106				
	(0.0128)	(0.0085)	(0.0078)				
First stage F-stat.	76	76	76				
Panel C. log(PM)							
$log(PM_{2.5})$	−0.0667	−0.0393	−0.0184	−0.0049	−2.255	−0.0264	−0.0161
	(0.0323)	(0.0133)	(0.0155)	(0.0024)	(0.8948)	(0.0142)	(0.0067)
First stage F-stat.	145	145	145	145	131	129	129
Observations	397,277	397,277	397,277	397,277	339,872	183,749	183,749
Panel D. $PM_{2.5}$ exceeding National Standard							
$PM_{2.5}$ > Standard	−0.1133	−0.0608	−0.0422	−0.0092	−4.536	−0.0379	−0.0244
	(0.0605)	(0.0281)	(0.0280)	(0.0043)	(2.0320)	(0.0230)	(0.0117)
First stage F-stat.	81	81	81	81	75	78	78
Observations	355,925	355,925	355,925	355,925	303,211	162,740	162,740

Notes: 2SLS estimates of the parameter β in Equation (1). In Panel A, the regressor of interest is $PM_{2.5}$ concentration measured in $\mu g/m^3$. In Panel B, an indicator for a $PM_{2.5}$ shock is used instead (as defined in Equation 3). In Panel C, the regressor is the logarithm of $PM_{2.5}$ concentration. In Panel D, the regressor is a binary indicator for $PM_{2.5}$ concentration violating the national air quality standard. Inverse hyperbolic sine transformations are applied to outcomes in Panels A and B. The first stage specification is given in Equation (2). Covariates as described in Table 2. Regressions are weighted by the number of active workers in a city during the current month. Standard errors clustered at the city level are reported in parentheses.

B Additional Figures

Figure 1: Skill Requirements in High-Skill Occupations and Software Development



Notes: Based on data from O*NET Database Version 25.0. Skills Table. Light blue bars reflect average importance of the respective skill across all high-skill occupations, i.e. occupations in Job Zones 4 and 5. Dark blue bars reflect the importance of the respective skill among software developers, i.e. occupation 15-1132.00 ("Software Developers, Applications").

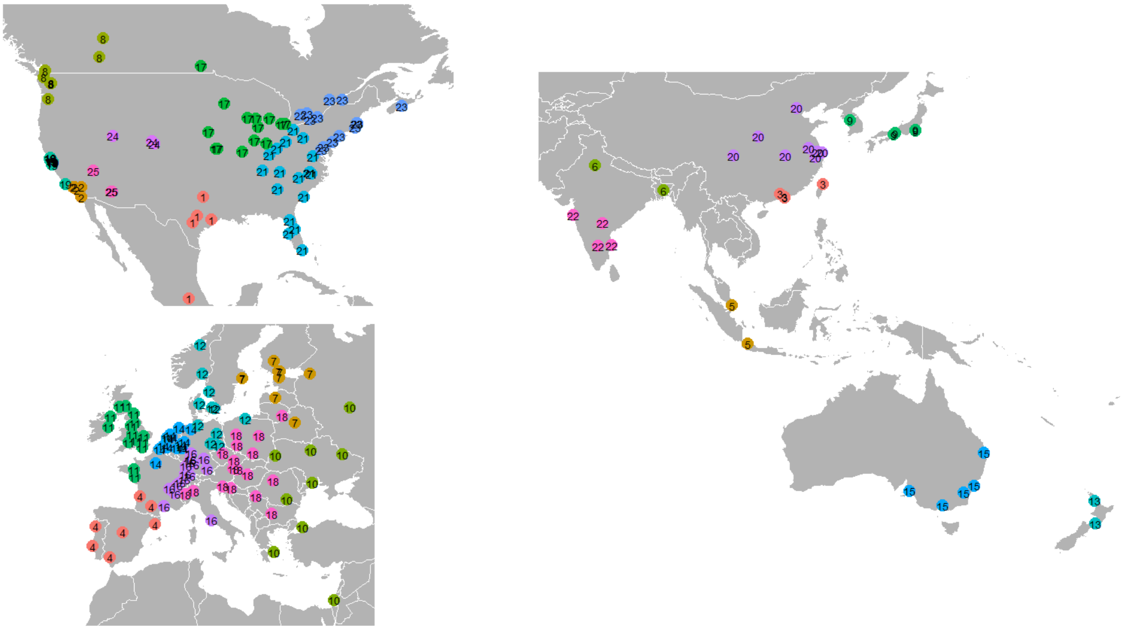
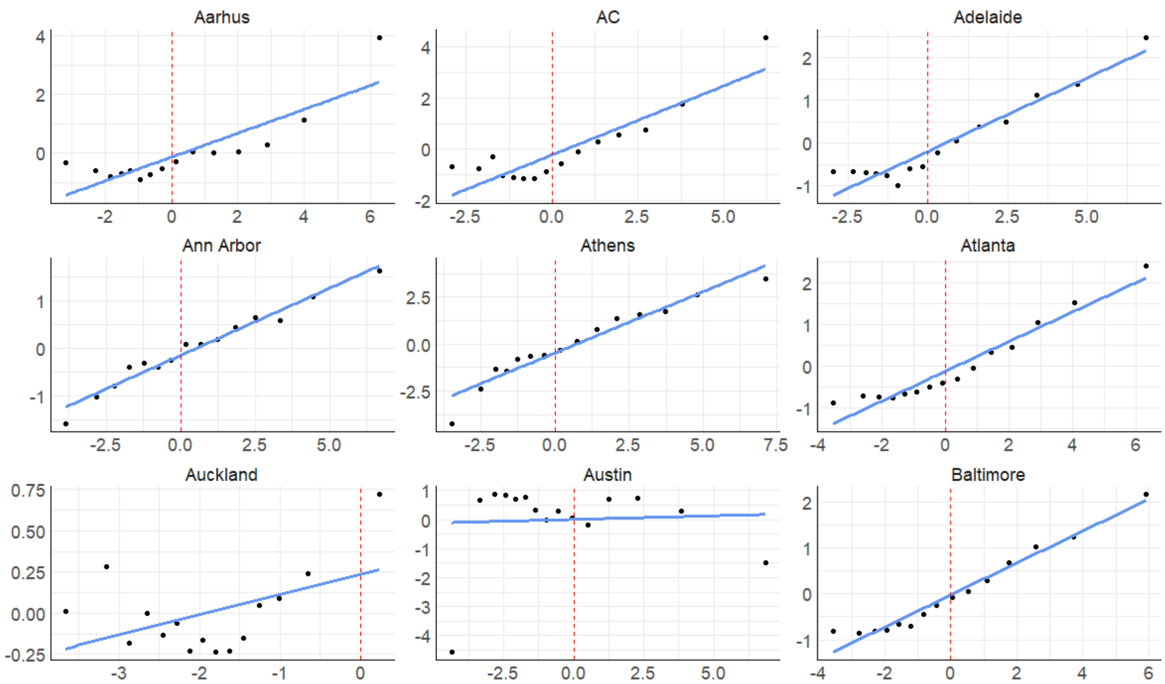


Figure 2: Illustration of first stage city groups g. Circles represent sample cities. Color and number of the city markers refer to the group g we assign a city to for the first stage regression (see Section 4 and Equation (2)).



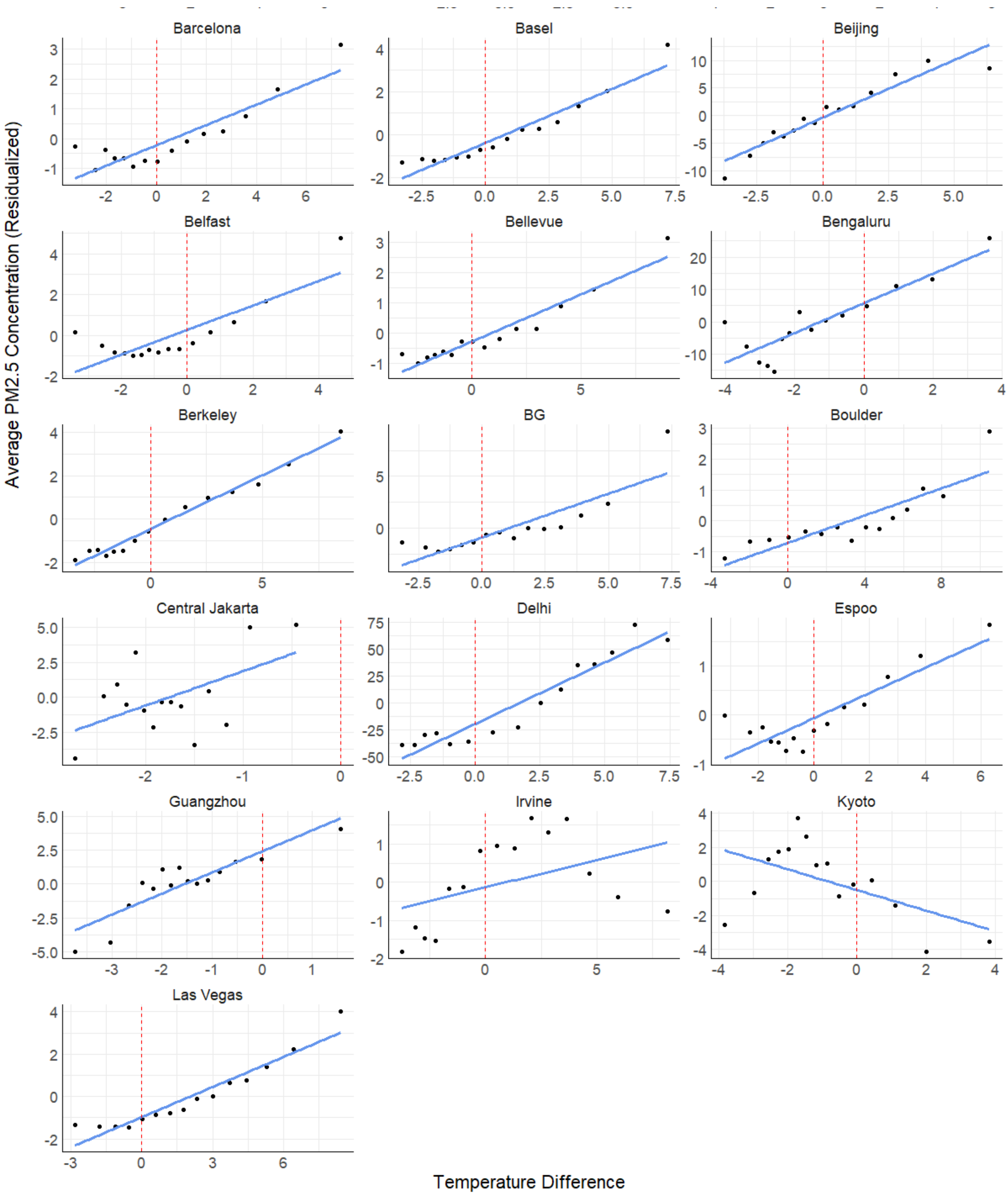


Figure 3: First stage for all 25 city groups. Binned scatter plots of residualized PM2.5 concentration against inversion strength for each first stage city group. Residuals are generated by taking out weather controls and fixed effects as described in Equation 2. Blue lines show a linear fit. Plot titles denote one city from the respective group.

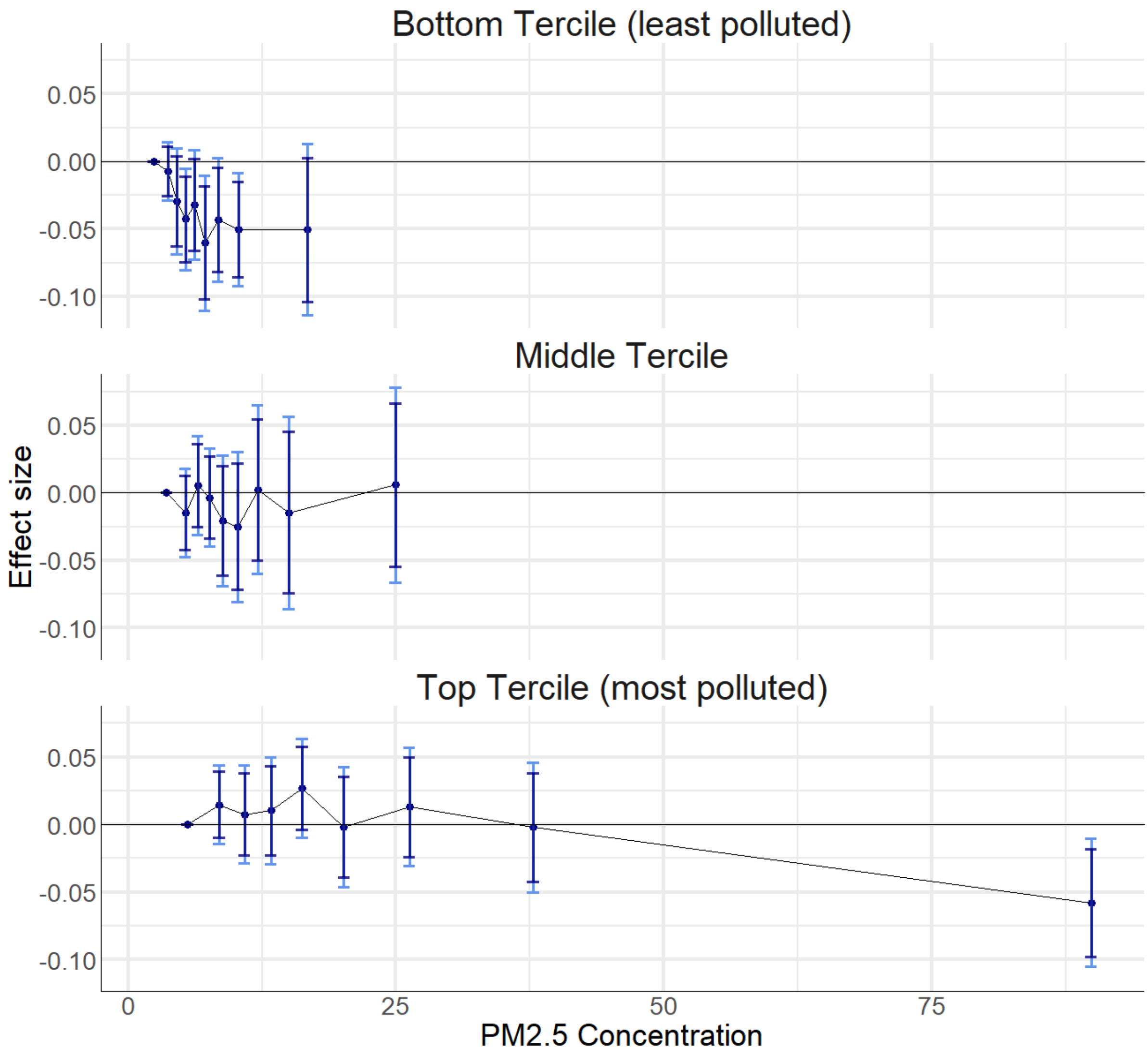


Figure 4: Non-linear effects of $PM_{2.5}$ on Work Quantity across subsamples based on average $PM_{2.5}$

Notes: The figure depicts point estimates on different bins of $PM_{2.5}$ concentrations from OLS regressions of total actions on indicators for each bin for three distinct samples. Cities are assigned to subsamples based on average $PM_{2.5}$ concentration. Covariates: Weather and holiday controls as in Equation 1, region × date and city × month fixed effects. X-axis: Average $PM_{2.5}$ concentration in each bin in µg/m^3. Error bars indicate 95%- and 90%-confidence intervals.

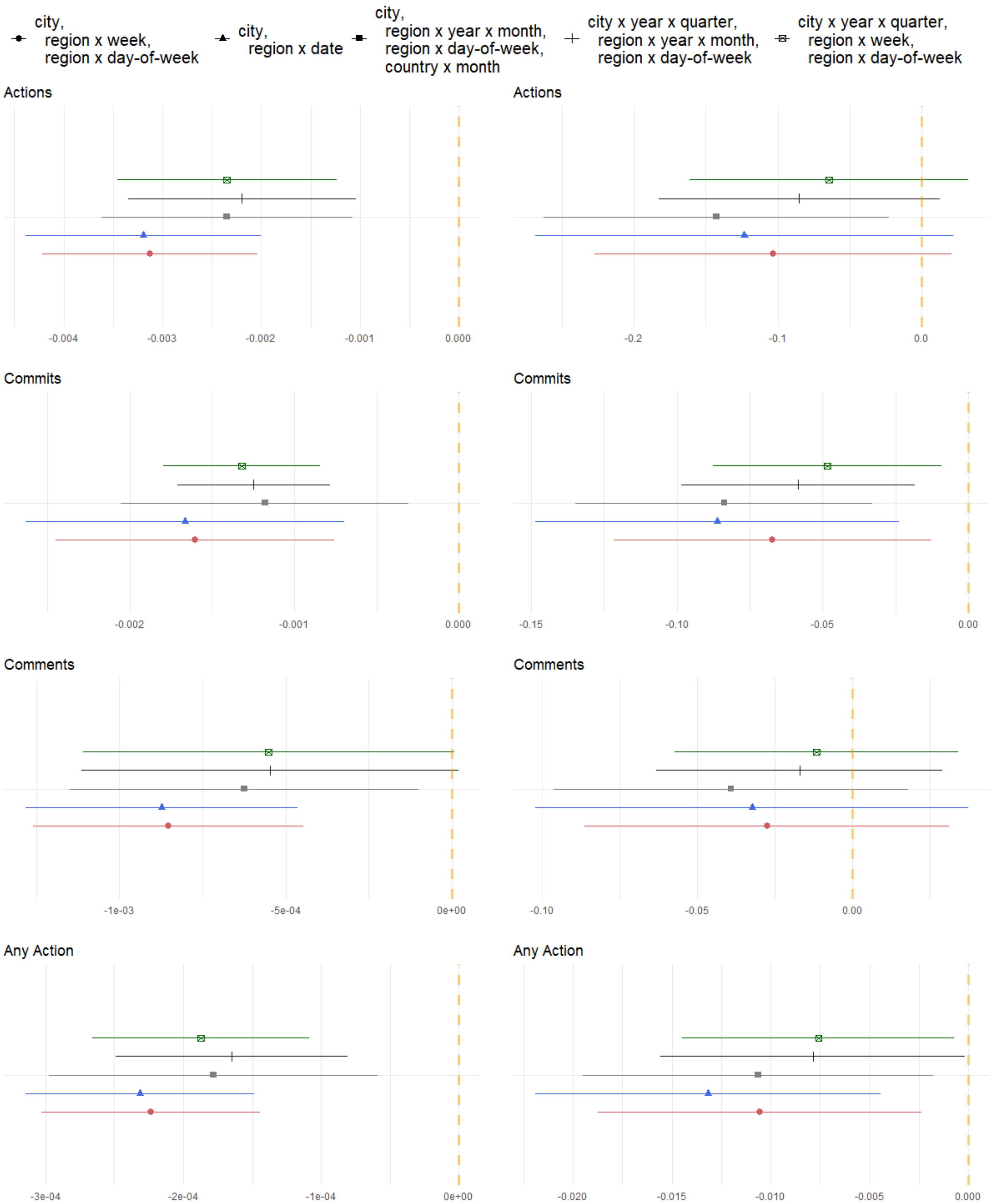


Figure 5: Robustness to Changes in Fixed Effects (Output Quantity)

Notes: IV point estimates of the parameter β in Equation (1), along with 95% confidence intervals. Dependent variables are denoted at the top of each plot. Left side: regressor of interest is $PM_{2.5}$ concentration in $\mu g/m^3$. Right side: regressor of interest is an indicator for a $PM_{2.5}$ shock. Relative to specifications underlying results in Table 2, we change the included fixed effects, as stated in the legend at the top. All regressions include control variables as described in Table 2 and are weighted by the number of active workers in a city during the current month. First Stage F-Statistics range from 67 to 207.

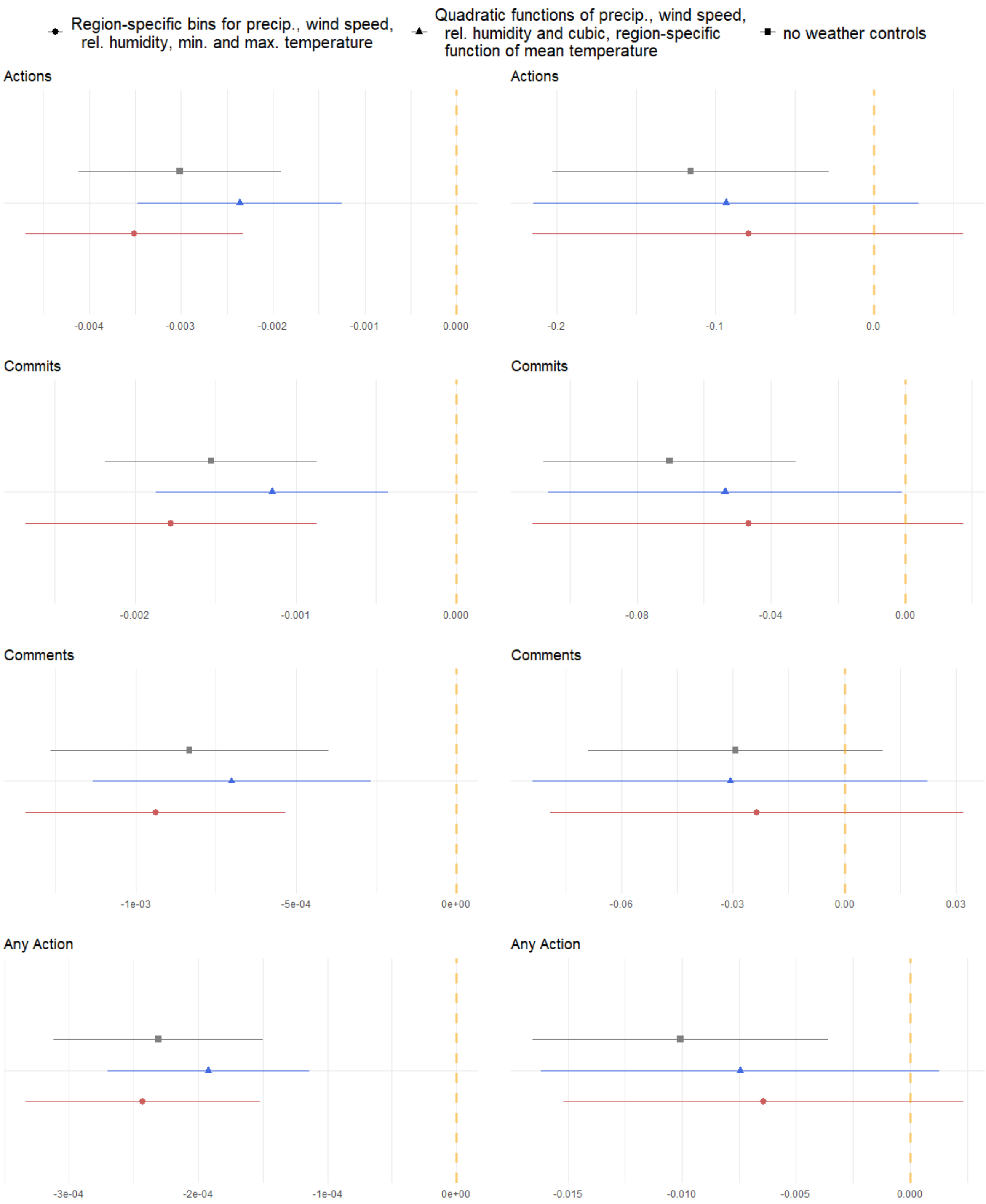


Figure 6: Robustness to Changes in Weather Controls (Output Quantity)

Notes: IV point estimates of the parameter β in Equation (1), along with 95% confidence intervals. Dependent variables are denoted at the top of each plot. Left side: regressor of interest is $PM_{2.5}$ concentration in $\mu g/m^3$. Right side: regressor of interest is an indicator for a $PM_{2.5}$ shock. Relative to specifications underlying results in Table 2, we change the included covariates to control for weather conditions, as stated in the legend at the top. All regressions include fixed effects as described in Table 2 and are weighted by the number of active workers in a city during the current month. First Stage F-Statistics range from 54 to 397.

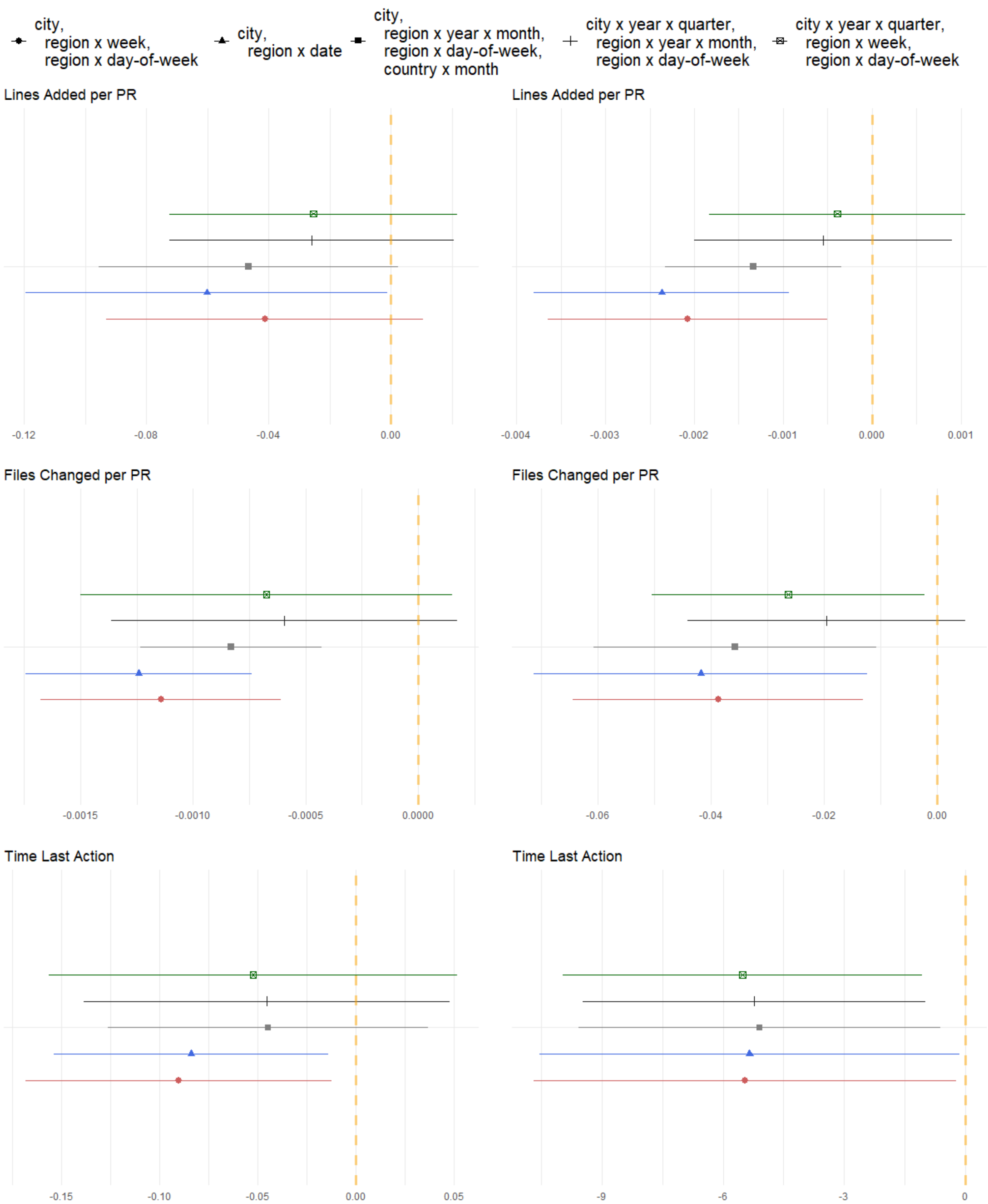


Figure 7: Robustness to Changes in Fixed Effects (Adjustment)

Notes: IV point estimates of the parameter β in Equation (1), along with 95% confidence intervals. Dependent variables are denoted at the top of each plot. Left side: regressor of interest is $PM_{2.5}$ concentration in $\mu g/m^3$. Right side: regressor of interest is an indicator for a $PM_{2.5}$ shock. Relative to specifications underlying results in Table 2, we change the included fixed effects, as stated in the legend at the top. All regressions include control variables as described in Table 2 and are weighted by the number of active workers in a city during the current month. First Stage F-Statistics range from 58 to 344.

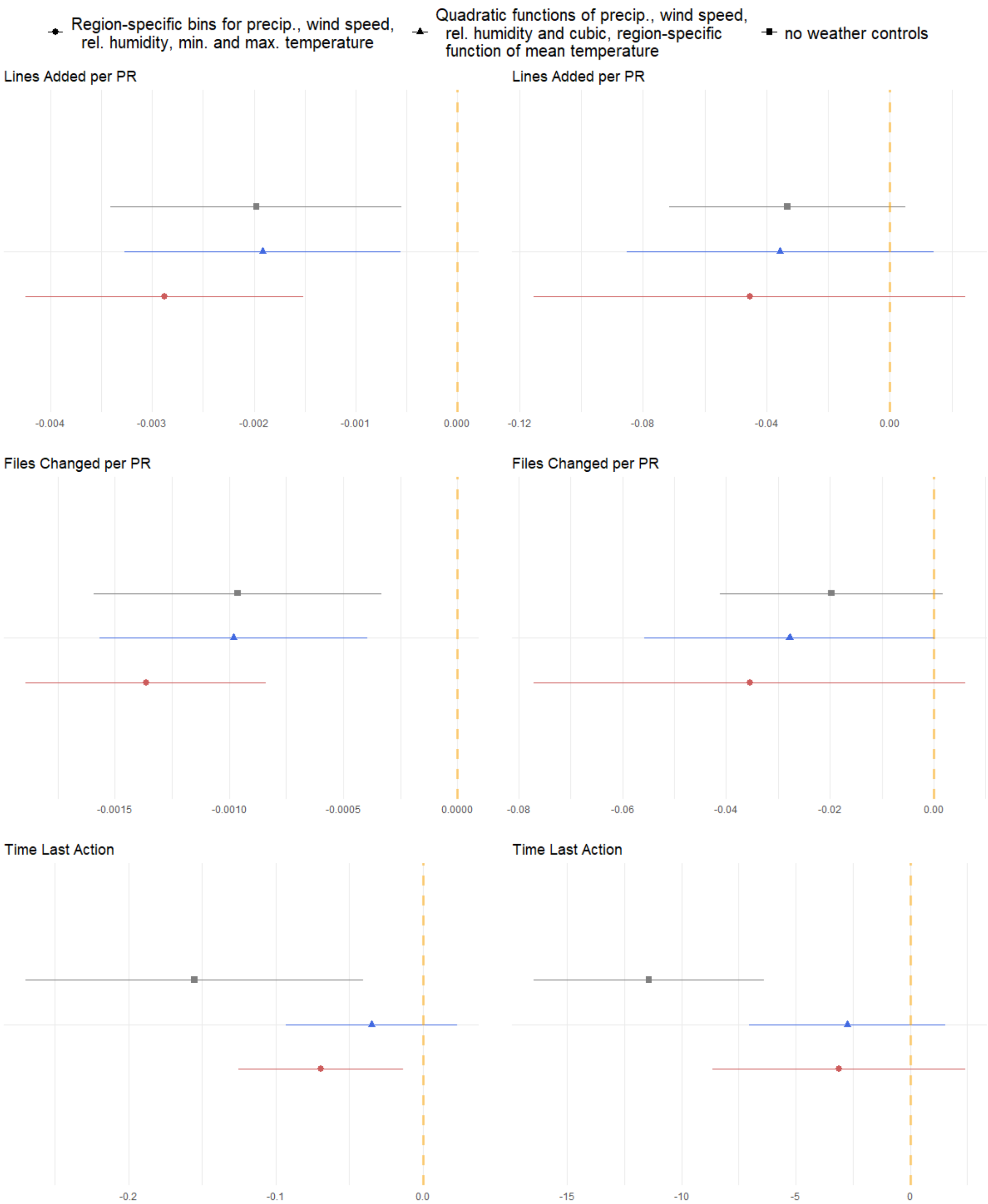


Figure 8: Robustness to Changes in Weather Controls (Adjustment)

Notes: IV point estimates of the parameter β in Equation (1), along with 95% confidence intervals. Dependent variables are denoted at the top of each plot. Left side: regressor of interest is $PM_{2.5}$ concentration measured in $\mu g/m^3$. Right side: regressor of interest is an indicator for a $PM_{2.5}$ shock. Relative to specifications underlying results in Table 2, we change the included covariates to control for weather conditions. We state the included variables in the legend at the top. First stage specification is given in Equation (2). All regressions include fixed effects as described in Table 2 and are weighted by the number of active workers in a city during the current month. First Stage F-Statistics range from 49 to 523.

C Gitcoin

This section provides additional details regarding the data collected from Gitcoin to assess the monetary value of output produced on GitHub and to validate some of our productivity and task complexity outcomes.

We collect data on 292 Gitcoin transactions via the Gitcoin API, including the type of the posted issue (bug, documentation, improvement, feature, or other), the expected issue difficulty as assessed by the issue funders (beginner, intermediate, or advanced), the URL to the PR solving the issue and awarded the payment, the value of the payment in USD, and the number of hours worked on the PR as stated by the PR author. The number of issues is relatively low compared to the volume of our GitHub data because Gitcoin is much younger than GitHub and only used by a small share of GitHub users. Using the URL of the PR, we combine this with information on pull request size obtained via the GitHub API, i.e. the number of commits it comprises, the number of lines of code added and deleted, and the number of files changed. This is possible because all Gitcoin issues and PRs are created in public GitHub repos and thus visible to us. In this context, a pull request reflects the complete work on a certain issue. Commits can be interpreted as single work steps in completing this task.

Combining the data on the amount of coding work done and on the payment made we can estimate the monetary value of output produced in public GitHub repos. The average monetary value per commit ranges from \$32 in the subsample of issues of difficulty level *beginner* to \$679 among issues marked as *advanced*. In the full sample, it amounts to \$112. The mean time input per commit also exhibits a steep gradient with respect to difficulty: It is 1 hour at the beginner level, but 5.3 hours at the advanced level.

To validate the use of the number of commits per day as one of our core measures of developer productivity, we analyze how the number of commits in a PR correlates with the payment awarded and the time spent on it in the Gitcoin sample.

Table 1 depicts results from regressions of the payment awarded for a PR, $log(payment_i)$, on the number of commits it comprises, $commits_i$ (columns 1-3), or the logarithm thereof (columns 4-6). We run specifications without any controls (columns 1 and 4), with controls for issue difficulty, issue type and the year of PR creation (columns 2 and 5), and alternatively with repository fixed effects (columns 3 and 6). The omitted difficulty category is *advanced*. Across specifications we find statistically significant positive effects, indicating that a higher number of commits is associated with higher payments. In terms of magnitude, the results from the regressions without any controls imply that one additional commit is associated with a 5.4% increase in payment (column 1), or that a 10% increase in the number of commits is correlated with a 3.5% rise in payment (column 4). When adding controls for issue difficulty and type, the magnitude of the effect is reduced. This reduction implies that part of the increase in payments in commits is driven by higher issue complexity. Even when using only variation across PRs submitted to the same repo, i.e., work on the same project, the positive relationship persists.

Table 1: Validity Check: Number of Commits and Gitcoin Payments

	$\log(payment_i)$					
	(1)	(2)	(3)	(4)	(5)	(6)
$commits_i$	0.054	0.039	0.034			
	(.010)	(.009)	(.010)			
$\log(commits_i)$				0.348	0.264	0.192
				(.071)	(.068)	(.059)
$\mathbf{1}\{Difficulty_i = Beginner\}$		−2.399			−2.412	
		(.439)			(.419)	
$\mathbf{1}\{Difficulty_i = Intermediate\}$		−1.878			−1.851	
		(.415)			(.405)	
Year dummies		✓	✓		✓	✓
Issue difficulty		✓			✓	
Issue type		✓			✓	
Repository FE			✓			✓
Observations	292	274	292	292	274	292

Notes: The table presents results from OLS regressions using data on the sample of Gitcoin pull requests. Observations are at the pull request level. Dependent variable is the logarithm of the payment awarded to the PR author. Explanatory variables are the number of commits (columns 1 to 3) or the logarithm thereof (columns 4 to 6). Columns 2 and 5 add dummies for the year the pull request was created, dummies for issue difficulty, and dummies for issue type. Column 3 and 6 instead add dummies for the year the pull request was created and fixed effects for the repository. Robust standard errors are reported in parentheses.

In Table 2 we present results from models where the dependent variable is $hoursworked_i$, the time input as reported by the PR author. We find that the time required to complete a task increases in the number of commits, and more so for issues of higher difficulty.

To validate our proxies for PR complexity, we run the specifications from columns 4 to 6 of Table 1 again, but add the number of files changed in the PR and the logarithm of lines of code added as additional regressors. Results are presented in Table 3. Holding the number of commits constant, adding more lines of code and changing more files is associated with a higher payment, suggesting that these variables indeed reflect task complexity.

Table 2: Validity Check: Number of Commits and Hours Worked on a PR

	$hoursworked_i$			
	(1)	(2)	(3)	(4)
$commits_i$	0.375	0.939		
	(.132)	(.341)		
$\log(commits_i)$			2.375	10.346
			(.648)	(3.535)
$\times\ \mathbf{1}\{Difficulty_i = Beginner\}$		−0.882		−9.748
		(.354)		(3.574)
$\times\ \mathbf{1}\{Difficulty_i = Intermediate\}$		−0.667		−8.471
		(.349)		(3.560)
Observations	271	267	271	267

Notes: The table presents results from OLS regressions using data on the sample of Gitcoin pull requests. Observations are at the pull request level. Dependent variable is the number of hours worked reported by the PR author. In colunm 1 the only explanatory variable is the number of commits in the PR. Column 2 adds dummies for issue difficulty and interactions between the number of commits and the difficulty dummies. The ommited difficulty category is *advanced.* In columns 3 and 4 report results from the same models except that the number of commits is replaced by the logarithm thereof. Robust standard errors are reported in parentheses.

Table 3: Validity check: PR complexity and Gitcoin payments

	$\log(payment_i)$			
	(1)	(2)	(3)	(4)
$\log(commits_i)$	0.143	0.136	0.070	0.145
	(0.067)	(0.068)	(0.058)	(0.056)
$fileschanged_i$	0.005	0.007	0.011	0.004
	(0.005)	(0.004)	(0.004)	(0.004)
$\log(linesadded_i)$	0.152	0.112	0.091	0.150
	(0.036)	(0.035)	(0.028)	(0.033)
$easylabel_i$				−0.348
				(0.173)
Year dummies	✓	✓	✓	✓
Issue difficulty dummies		✓		
Issue type dummies		✓		
Repository FE			✓	
Observations	292	274	292	270

Notes: The table presents results from OLS regressions using data on the sample of Gitcoin pull requests. Observations are at the pull request level. Dependent variable is the logarithm of the payment awarded to the PR author. Explanatory variables are the number of commits and the number of lines of code added in the PR (both in logs), the number of code files changed and dummies for the year the pull request was created. Column 2 adds dummies for issue difficulty and issue type. Column 3 instead adds fixed effects for the repository. Column 4 instead adds a dummy variable taking a value of one if the issue addressed by the PR carries a label that we classify as indicating an easy issue. The number of lines of code added and of files changed in the PR are winsorized at the 1st and the 99th percentile. Robust standard errors are reported in parentheses.

D Auxiliary Regressions

For estimating equation (1), measures of the output of each individual developer i are aggregated to the city-day level. Instead of forming simple averages, we take into account additional information at the developer level. This is done by estimating auxiliary regression, a common approach in this literature (e.g. Currie et al., 2015). In a first step, we estimate regressions for outcome y of developer i living in city c on day d of the following kind.

$$y_{i,c,d} = \mu_i + \mathbf{x}'_{i,d}\pi + \psi_{c,d} + \varepsilon_{i,c,d} \tag{1}$$

Here, $y_{i,c,d}$ denotes one of the measures of developer output, task choice, or working hours. The fixed effect μ_i captures time-invariant unobserved factors at the developer level. Including these is important as the composition of developers changes over time. A developer's experience is controlled for by $\mathbf{x}_{i,t}$, a vector of indicators for time since registration on GitHub, where each indicator represents a time span of three months. Additionally, equation (1) includes city-day fixed effects. Their estimates $\widehat{\psi}_{c,d}$ give the average outcome for a city-day after controlling for experience and composition effects. These estimates replace the dependent variable in equation (1).

This approach is computationally less costly and asymptotically equivalent to directly estimating the regressions at the individual developer level (Donald and Lang, 2007). We take into account the sample variance by weighting all regressions by the number of underlying developer observations in each city-day cell (cf. Currie and Neidell, 2005; Isen et al., 2017).