How Does Improvement in Commuting Affect Employees? Evidence from a Natural Experiment

Yao Lu, Xinzeng Shi, Jagadeesh Sivadasan, and Zhufeng Xu*

Abstract

We collect worker month-level panel data from two companies for a two-year period before and after the opening of a nearby subway station, which significantly improved public transportation commutes for some workers. We find a significant difference-in-differences increase (12.6% of the standard deviation) in bonus pay (which is strongly correlated to worker-
level performance measures) for affected workers relative to unaffected coworkers. We find no evidence that the improved performance is a result of affected workers spending extra time at the workplace. We find suggestive evidence for a relative decline in turnover, consistent with a gain in utility for affected workers.

**Keywords**: labor productivity, subway, transportation, worker turnover, incentives

**JEL Codes**: H54, J63, J24, J22

1. Introduction

Numerous studies document that commuting is a highly disliked activity, and that tougher (i.e., longer, more crowded, and/or more unpredictable) commutes are associated with lower subjective well-being. ¹ To improve commutes, governments continue to make huge investments in transportation infrastructure, particularly for urban transit (Redding and Turner, 2015; Lu et al., 2016; Freemark, 2018; Koyanagi, 2017).

Because infrastructure for commute improvement is costly, it is important to understand how such improvements affect economic outcomes. However, because the location of transportation infrastructure investment is often influenced by local economic conditions, it is hard to identify causal effects (Redding and Turner, 2015). While a number of important studies have used alternative identification approaches to assess the impact of transportation infrastructure on aggregate region-level outcomes, studies examining worker-level outcomes

¹ Kahneman et al. (2004) finds that respondents view commuting among the least pleasant of daily activities. Studies documenting association of longer commutes with lower measures of subjective well-being include Choi, Coughlin, and D’Ambrosio (2013) for US, Chatterjee et al. (2017) for UK, and for China.
are rare, limiting our understanding of how workers benefit from infrastructure investment aimed at improving their commutes.2

We contribute to the literature by obtaining and exploiting rich worker-month level “insider” panel data, in a natural experiment setting that allows us to examine how workers and firms responded to an improvement in public transit. In particular, we examine the effects of the opening (in December 2014) of the terminal station of Subway Line 15, which is located underneath a plaza at the center of a busy business district in Beijing, China (see Figure 1; Appendix OB1 provides more background information). Subway Line 15 significantly improves the commute quality for affected employees in several ways. First, the availability and proximity of the new subway line reduced the public transport round-trip commute time substantially, by a median amount of 25.3 minutes (or 5.3% of an 8-hour workday) for workers able to use the subway. Second, the subway commute is much more predictable relative to other modes involving road traffic, which have uncertain congestion delays. Finally, because the new station is a terminus, the commute on this subway is significantly less crowded (especially for return commutes), reducing an important commute-related stress factor (Singer, Lundberg, and Frankenhauser 1978) and making it easier to work during the commute.

Our data and setting allow us to adopt a difference-in-differences (DID) estimation strategy. Specifically, we compare the before-after change of outcomes within employees who can take

---

2 Baum-Snow (2007), Duranton and Turner (2011) and Donaldson (2018) exploit variation induced by historical planned routes. A number of papers have examined aggregate regional outcomes from Chinese infrastructure investments (e.g., Banerjee, Duflo and Qian 2020; Faber 2014; Baum-Snow et al. 2017). Gu et al. (2019) document that subways alleviated road congestion in China.
advantage of the subway line (i.e., those for whom the fastest public transport route to work involves Subway Line 15, hereafter called “affected” workers) to the within-employee changes for those who do not benefit directly from Subway Line 15 (hereafter called “unaffected” or “control” workers).

We focus on two primary worker-level outcomes: worker performance pay (bonus) and worker retention. Indeed, the bonus pay of affected employees shows a significant DID increase of about 12.6% of the standard deviation, following the opening of the new subway line. Our analysis of bonus and available (post-subway opening) data on worker-level performance scores, as well as our discussions with management, confirm that bonus pay is strongly linked to worker-level performance. Hence, we interpret the baseline results as implying that better commute led to better worker performance.3 The baseline result is robust to using a propensity score matched control group where we get better balance on observables between the treatment and control groups. Event study analyses confirm that prior trends were parallel for affected and unaffected employees, and reveal a significant relative increase in bonus pay for affected workers coincident with the opening of the new subway station.

We verify that the increase in bonus for affected workers is robust across a number of other checks. Specifications without individual fixed effects and using a balanced panel sample,

3 We are careful not to interpret our results as reflecting an increase in worker productivity (i.e., greater output from same time consumption), as we cannot rule out the workers spending some of the saved commute time, either on the commute or at home, to generate the performance improvement. While our results are consistent with the subway serving as a productivity-enhancing technology that reduces “dead time” wasted in commutes, we use the broader worker “performance” term throughout.
suggest some modest positive selection effects (from differential retention of affected workers). Utilizing individual-level annual sales data available for marketing personnel, we directly confirm that the bonus results reflect improvement in underlying worker performance. We verify that bonus increases for affected workers do not negatively impact subordinates or coworkers. We find negative bonus elasticity to commute time, confirming that bonus increases were larger for affected workers with greater commute time savings.

Assuming that worker performance is a product of worker intensity of effort (i.e., output per unit of time or productivity) and quantity of work time suggests these two plausible channels for our results. Surprisingly, we find no evidence for an increase in time spent at the workplace, or for any effect on absenteeism, or for affected workers spending additional time at the office. Workers we interviewed did not think the subway has affected punctuality or time at work (as they had already been highly motivated by company mandates before the opening of the subway). These findings suggest a prominent role for increased intensity of effort, potentially fueled by reduced fatigue and increased well-being.4

Next, we investigate the effect of the new subway station on employee turnover. Both regression estimates and event study graphs suggest that turnover declines for affected workers

---

4 A rich literature finds that shorter, less crowded and more predictable commutes are associated with higher subjective well-being (e.g., Chatterjee et al., 2017; Zhu et al., 2017), and a separate literature has documented strong links between workers’ happiness and workers’ effort and creativity (Porto, 2016). Affected interviewees confirm that they value the improved “predictability of the commute”, feel more energetic, and are able to devote more time to exercise, sleep, and caring for family members. Improvement in sleep and exercise could also play a role in improving worker effort productivity (Lechner, 2015; Gibson and Shrader, 2018).
relative to unaffected workers. We also find some evidence for an increase in quality of hires (conditional on wage) from affected locations. However, we do not find significant relocation by workers towards the subway, or significant new hiring of subway users.

A broader event study analysis of firm market valuations using 13 subway station construction commencement events across Beijing over a longer time period shows evidence for positive abnormal returns around the construction commencement events. This analysis provides supplementary evidence that the benefits accrue to shareholders when their firms gain access to a new subway, and that our findings potentially hold for a broader sample (assuming that worker performance is a key driver of firm value).

Finally, we undertake a number of assessments of aggregate surplus from the new subway. Combining estimates of the effect of the subway on individual-level sales and on a measure of worker welfare (which captures disutility from commute valued at half the wage rate per Redding and Turner, 2015), we get an estimate of the elasticity of worker welfare to sales of 0.065, which is within the 0.05 to 0.15 range documented in the rent-sharing literature (Card et al. 2018). A back-of-the-envelope estimation (based on a number of strong assumptions) of benefits to affected workers within a radius of 500 meters from the new subway station suggests an annual income gain of 56.6 million RMB (or about 4% of the estimated construction cost of about 1.4 billion RMB), and annual commute savings benefit of 51.5 million RMB (or about 3.6% of the construction cost of the new subway station).

In addition to presenting novel evidence for a positive impact of commute enhancement on worker performance and retention, our paper contributes to three related streams of research. First, it extends the recent literature studying the effects of workplace and external
environments on worker performance. Close in spirit to our work is an influential study of
telecommuting by Bloom et al. (2015) that tackles identification issues by randomizing workers
into telecommuting, and finds that working from home significantly improves worker
performance. While their results are consistent with ours in that elimination of commute is
associated with improvement in worker performance, telecommuting changes other important
aspects of the employment relationship, including the degree of monitoring and visibility of
telecommuting workers, and social interactions between workers, coworkers, and managers.

Second, our paper contributes to the literature examining the impact of commuting on
individual-level outcomes. Fu and Viard (2019) find a reduction in work time for faculty from
commute increase, and van Ommeren and Gutierrez-i-Puigarnau (2011) find that commuting
distance is positively correlated with absenteeism; we find no direct impact on punctuality or
time at work. Monte, Redding and Rossi-Hansberg (2018) present a model (and estimates) for
welfare increase from reduction in commute costs, which allows workers to reside in high
amenity locations and commute to work in high productivity locations; we contribute to the
related economic geography literature by highlighting how commute reduction can generate
welfare gains by improving worker performance.

Third, our paper contributes to the broad literature investigating the drivers of firm-level
productivity (Syverson, 2011), and specifically complements studies finding that better traffic
conditions increase firm and aggregate productivity (e.g., Tang, 2014; Firth, 2017). Our paper
finds that access to subway stations can improve employee-level performance, which could be
an important source of improvement in firm productivity.
2. Brief Theoretical Motivation

In this section, we discuss the expected effects from the opening of Subway Line 15. (Two versions of a formal principal-agent model are presented in the Online Appendix OA). We expect the reduction in commute time to increase effort by workers for three related reasons: (i) a reduction in fatigue which lowers the cost of effort; (ii) an increase in worker wellbeing which increases worker effort (Porto, 2016); and (iii) reduced total time away from home, which allows for more focus while at work, as workers have more time at home to complete errands and other non-work activities. Consequently, we expect the productivity of affected workers to increase after the opening of the subway, so monthly individual-level bonus payments that are linked to measures of worker performance should go up for affected workers. Shorter commutes provide a non-pecuniary benefit, which the firm could potentially claw back through lower base pay; thus the impact on total compensation for affected workers is ambiguous.

Further, we expect the increase in non-pecuniary utility from better commute to reduce the likelihood of affected workers exiting the firm (Gronberg and Reed 1994), assuming that the completion of the terminal station does not have a large impact on the set of outside offers. We expect the commute amenity to also allow the firm to attract better workers (for the same level of compensation) from applicants who are able to use the subway.

3. Data, Company Backgrounds, and Variables

3.1. Data and Company Backgrounds

Our quantitative data comes from two sources. We obtained administrative data on past and current employees from the two companies, and we also conducted a survey to collect other
information from current employees of these two companies. The two companies are subsidiaries of a common parent company, which cooperated with us for this study. Company One was founded in 2000. Its business activities include software services (computer system service, system integration, and data processing) and hardware sales (computers, software and auxiliary equipment, mechanical and communication equipment, and electronic products). Company Two was founded in 2006. It is a leading provider of smart energy savings services in China. Its business covers seven major fields: intelligent transportation, intelligent building and community design and management, building construction energy saving, intelligent energy, building automation systems, fire alarm systems, and urban security. Company One (Two) had revenue of 1,732 Mn RMB (1,786 Mn RMB) in 2016. Company One (Two) had 260 (110) employees in December 2014. Analysis using national tax survey data suggest these companies are similar to medium-sized firms in the IT sector in China, and similar in employment size to other firms located in Tongfang Plaza (Appendix OB2). We also draw on two rounds of interviews with employees of both firms for qualitative evidence.

A. Administrative Data

This data set, obtained from the two companies, includes detailed information related to fixed (time invariant) characteristics, as well as time-varying monthly data, for every employee who ever worked in these two companies between 2013 and 2016. The main time-invariant variables we use include joining date (which helps measure tenure at the firm), exit date (if the employee left the company before the end of our sample period), and the date of start of the first job (which helps measure total experience). This data set also provides information on gender, birthday, ethnicity, hukou type (whether household registration is agricultural or non-
agricultural), and education (including college, major, degree, and gradation year).

The two main components of employee monthly compensation are base salary and bonus. Monthly information on each employee’s attendance includes the number of days that the employee arrived late, left early, or took personal leave or sick leave.

In our baseline analysis we drop a small number (16, or 4% of total) employees that moved homes to retain only on employees who did not relocate after Line 15 opened, and we include only those employees who are in the company workforce for at least one month before and after the opening of the subway. Our sample includes 265 (112) employees in Company One (Two), of which 105 (39) are affected (Appendix Tables OB1a and OB1b provide more details).

B. Survey Data

To complement the administrative data, in July 2017 we conducted a survey of all employees in the two firms. We use the responses for three types of information: (i) changes of home addresses since they joined the company; (ii) self-reports of use of, and time saved by, the subway; (iii) current and past (recalled) work-life balance for each year of 2013-2017. We link the information in the survey data with the administrative data using employees’ work ID. The survey response rate was about 85% of surviving employees in 2017.

3.2. Definition of Affected Workers and Related Variables

The primary focus in our analysis is to estimate the analog of “intent-to-treat” effects, where the “treatment” or “affected” group comprise those workers whose public commutes are shortened by the new Subway Line 15. This provides a conservative estimate (in the presence of some affected workers who may choose not to take the subway) and is a relevant parameter of interest for policymakers (Duflo, Glennerster and Kremer 2008). Specifically, to construct
this “affected” worker dummy variable, we first use Baidu Direction API to obtain all public transportation routes as well as associated travel time from the employee’s home address to the company. We then define affected workers (denoted by a NearSubway dummy variable) as those for whom the fastest public transportation route from his/her home address to the office includes Line 15. Overall, 38.2% of the employees in our cleaned sample are affected by the subway (see row 4 of Col 1 in Panel A of Appendix Table OB1a). As an alternative benchmark, we also examine effects on the workers who self-report (in our worker survey) to be users of the new subway; using this dummy yields a “treatment on the treated” (or per-protocol) effect as it compares workers who actually use the subway to controls that do not. The online Data Appendix Table D1 provides a more detailed description and analysis of the NearSubway dummies based on both Baidu and self-reported data.

3.3. Key Outcome Variables

A. Summary Statistics

Table 1 reports the summary statistics for the main variables in the baseline sample. Bonus and income variables are measured monthly and deflated to 2013 using the CPI (Beijing Statistical Yearbook, 2017). For a small number of employee-month observations, we find that the bonus amount was negative in the data; we confirmed with the management that this reflects a genuine underperformance penalty. In order to retain the information, while undertaking analysis using logged variables we adjust all of the bonus amounts, as well as the total income, using a constant such that the minimum value of the adjusted bonus is equal to one. We confirm robustness to using alternative measures (Appendix Table OC10).

The mean of NearSubway dummy is 0.39, indicating that 39% of worker-month
observations relate to affected workers. The affected workers save, on average, about 21.8 minutes on a one-way commute. There are large outlier values in Baidu estimates for time saved, so the median time saved of 12.7 minutes one-way (about 25.4 minutes round trip) may be more representative of savings for the average worker. The average monthly bonus is 4,945 (2013, in RMB), which is about 41.96% of the average total income of 11,785 (2013, in RMB).

Examining demographic variables (Appendix Table OB2), we find affected workers are younger, more likely to be female, have less experience and tenure, and are less likely to be managers. This lack of balance raises potential identification concerns. To address this, we posit a model that predicts propensity to be affected by the subway, with distance from work, all of the demographic variables and a dummy for new employee entering the panel after January 2013 as explanatory variables (Appendix Table OB3), and then use the optimal matching approach in Ho et al. (2011) to construct a one-for-one matched sample. We confirm significant improvement in balance on observables, with no significant difference in the mean of the wage or demographics variables between the affected and control groups in the propensity-matched sample (Col 12, Appendix Table OB2).

B. Verifying the Worker-Level Bonus-Performance Link

Our interviews with managers and workers at both companies confirmed that bonuses are indeed linked to quantitative and qualitative evaluations of worker performance. To further confirm a direct link between bonus and worker-level performance measures, we regress the individual monthly log (bonus) on a one-year lead of the individual sales measure (available for marketing workers in both companies), as well as on performance scores available on a
monthly basis for Company One for 2015 and 2016. The results (Appendix Table OB5) confirm that variation in bonus is indeed strongly correlated with individual sales measures (in Panel A) and with worker performance scores (in Panel B), both in the cross-section across workers, and within-workers as well.

4. Analysis of Worker Bonus Pay

4.1. Empirical Strategy

The baseline DID regression specification is:

\[ y_{ict} = \beta \cdot NearSubway_i \times T_t + \alpha_i + \alpha_{ct} + \varepsilon_{ict}. \]  

Here, \( y_{ict} \) is the outcome variable for employee \( i \) in company \( c \) in year-month \( t \). \( NearSubway_i \) is a dummy variable equal to one for the affected employee \( i \). As discussed in Section 3.2, our primary interest is in the average “intention-to-treat” effect of the improvement in public transport commute and accordingly, we use a definition of “affected” workers (described in more detail in Section 3.2 and the online Data Appendix) that indicate workers whose public transport commutes were shortened by the subway. As an alternative benchmark, we define affected workers as those self-reporting (in our survey) as users of the subway. While this self-reported dummy captures “treatment-on-the-treated” (or per-protocol) effects which

\[ \text{We use a one-year lead of the sales measure to account for the fact that bonus is paid based to a significant extent on sales initiated in the current period (and other measures that capture current period worker effort), while the individual sales measure we have is sales executed in the current year (but initiated earlier). Per our discussion with management, performance measures were set as follows: the immediate superior assigns each employee a workload (a set of tasks) every month. If the employee finishes this set of tasks on time, her performance score would be 100 points; if she completed more (less) than her allocated tasks, the measure is adjusted proportionately.} \]
could be confounded by the self-selection of individuals more likely to benefit from the shorter commute, this effect is nevertheless of interest in evaluations. $T_t$ is a dummy variable equal to 1 for the period after the opening of Line 15 station and 0 otherwise. We control for individual fixed effects $\alpha_i$ to remove any individual-level time-invariant factors potentially correlated with the outcome variable, and use company-year-month fixed effects $\alpha_{ct}$ to allow period-specific shocks to vary across the companies. $\varepsilon_{ict}$ is the error term. To deal with correlation in the error term within individuals, as well as across individuals within a period, we estimate standard errors by allowing for two-way clustering over individuals and year-months. The coefficient of interest is $\beta$, which captures the DID change of the dependent variable following the opening of the subway station.

Because we compare employee-level outcomes within companies, there is less concerns about potential bias from firms lobbying for placement of the subway station. Other identification issues are addressed in Section 4.2, and additional robustness checks are presented in Section 4.3.

4.2. Baseline Results

Panel A of Table 2 presents baseline DID results (Equation 1). In Cols 1 to 4, we use Self-reported NearSubway constructed from self-reported use of the Subway, while in Cols 5 to 8, we use NearSubway constructed using Baidu Maps (see discussion in Section 4.1).

In Cols 1 and 5 we find an aggregate effect of about 0.03 (statistically significant in Col 5) of the opening of the subway on the log of total income. As discussed in Section 2, the impact on total income could be muted if firms claw back the benefit of lower commute by lowering the base pay; in Cols 2 and 6 we do find a negative (but noisy) effect on the log of the base
pay. The positive effect for total income suggests that the worker gains through the effort-bonus channel is not fully offset by claw back by employers through a reduced base pay.

The major focus of this paper is on worker performance reflected in the bonus pay; in Cols 3, 4, 7 and 8, we find a statistically significant effect on bonus. The larger “treatment on treated” estimates in Cols 3 and 4 of Panel A imply that being on the subway increases bonus by about 800 to 950 RMB per month, or about 17.0% of the standard deviation in bonus. The “intent to treat” estimates using the Baidu-based NearSubway status imply that taking the subway increases monthly bonus by about 640 to 580 RMB per month, or about 12% of the standard deviation in the bonus variable. The smaller “intent to treat” estimates are what we expect, as those within the affected group who do not take the Subway can be expected to have no performance effects. As discussed in Section 3.2, our primary interest is in the more conservative “intent-to-treat” estimate, which is the main focus of the remainder of the paper. Based on estimates of the linear relationship between performance measures and log bonus (Appendix Table OB5), and summary statistics on the performance score (Appendix Table OB4a), the estimated effect of the subway on bonus in Col 8 of Panel A is equivalent to an increase of about 0.104 points (or about 0.43 of the standard deviation) in the scaled performance score.

---

6 The lack of a decline in base pay (especially in Col 6 of Panel A) could be because cutbacks to non-bonus pay only for affected workers could be perceived as unfair (i.e., for no fault of the worker), or violate seniority and/or position norms.

7 The estimated effect of the subway in RMB and as % of the standard deviation (SD) of the dependent variable are provided in last two rows of Panels A and B. As an illustration, in Col 4, the estimated effect of 0.067 implies a change of \((0.067 \times (5076 + 6905)) = 803\) RMB in level terms (after accounting for the linear shift of 6905 in the adjusted bonus variable), or 17% \((= 0.067/0.39)\) of the standard deviation (from Table 1, Col 4) of the dependent variable.
performance score.

In Panel B, we undertake the same analysis using the propensity-score matched sample (discussed in Section 3.3.A). The results are similar to those of Panel A. Estimates for log total income are noisier; however, the estimated effect on the bonus measures are somewhat larger in Cols 3, 4, 7 and 8 of Panel B relative to Panel A. These results confirm that the different characteristics across the affected and unaffected workers (evident in Cols 5 to 8 of Table 1) do not significantly impact the baseline estimates of the effect of the subway on bonus. In Section 4.3.C, we confirm robustness to an alternative Coarsened Exact Matching (CEM) approach.

The DID estimation strategy implicitly assumes that the pre-existing time trends of outcome variables are parallel for affected and unaffected employees. To test this, we examine an event study graph (Figure 2) that presents the differential trend for affected workers (relative to unaffected) conditional on individual and company-quarter fixed effects, before and after the subway line opening. Figure 2 confirms that there is no differential time trend for the affected group of workers prior to the opening of the subway for our main variable of interest: bonus income. Moreover, immediately following the opening of the subway, there is a significant jump in the relative trend in bonus for the affected group, and this relative change remains persistently positive until the end of our sample period.8

4.3. Robustness Checks and Extensions of Baseline Results

We verified parallel trends for raw measures of bonus and total income excluding bonus before Subway Line 15 station opened, and a relative increase in bonus for the affected group after the opening of the subway, for the full sample (Appendix Figure OC2), a propensity-score matched sample (Appendix Figure OC3), and a balanced sample (Appendix Figure OC4).
We conduct a battery of robustness checks and extensions of the baseline results for our main variable of interest: worker performance proxied using log bonus variable as the preferred measure (most results and more details of the tests are in the Online Appendix OC).

**A. Within-Worker Changes vs Composition Effects**

To understand how firm-level aggregate effects are impacted by worker composition changes, we examine results when excluding individual fixed effects, using a fully balanced sample, and after including workers who moved after opening of the subway. The results (Appendix Table OC1) suggest a modest positive selection of high performing affected workers after the subway opening, but the differences are small. We also confirmed no systematic pre (or post) trend in the bonus-commute relationship (Appendix Figure OB3).

**B. Annual Marketing Sales**

One potential concern is that the bonus increase after the opening of the subway station is not from improvement in worker performance, but due to some non-performance-related reason such as because affected workers potentially get better outside offers, which prompts the companies to offer the affected workers higher pay (including higher bonus pay). We address this concern directly using an available measure of worker performance: individual-level annual sales for marketing personnel. We define the Post$_t$ period as equal to 1 for the post-2015 period (instead of 2014), to account for lags between executed sales (which the dependent variable captures) and contracted sales. The results (Table OC2) suggest large DID increase in sales for affected individuals, confirming that the pathway to higher bonus is indeed through greater worker output. We use available data on performance scores (for one of the
companies in 2015 and 2016) to confirm that affected workers were not rewarded more strongly for performance (Appendix Table OC7).

C. Other Robustness Checks

We conduct a number of other checks to further verify the robustness of our baseline findings - detailed explanations, including tables, are available in the Online Appendix OC. In particular, we: (1) apply permutation tests (re-testing across a 2000 randomizations of affected status) to verify significance to this alternative inference approach (Figure OC5); (2) use self-reported information from survey data to address potential measurement errors in Baidu-based measure of affected status (Table OC3); (3) drop employees moving before the opening of subway line (Table OC4); (4) control for the effects from housing costs which could be simultaneously impacted by the new Subway, to verify that changes in bonus pay were not the result of additional effort induced by increased housing costs for affected workers (Tables OC5 and OC6); (5) address concerns about the potential contamination effects from other subway lines (Table OC8) – we find that results are stronger, consistent with reduction of crowding and improvement of commute on a nearby line; (6) use alternative samples excluding controls close to affected workers (Table OC9) and bonus measurements with alternative ways to control for outliers (Table OC10); (7) use one-way clustering on individuals only (Table OC11); (8) use CEM (Iacus, King and Porro, 2012, 2020) as an alternative to propensity score matching (Table OC12a and Table OC12b). We find our baseline results are robust to each of these checks; (9) We verify that spatial characteristics likely to be correlated with affected status and spatial shocks do not explain the observed baseline effects (Table OC13). (10) We explored heterogeneous effects by position and occupation. We find stronger bonus effects for marketing
workers relative to administration and technology workers, and non-managers relative to managers (Appendix OC11 and Appendix Table OC14).

D. Do Effects Come at the Cost of Subordinates and Co-workers?

An important potential concern is whether there are negative spillover effects to peers and subordinates, as e.g., could arise if the aggregate bonus pool was fixed, and/or relative performance measurement was used. The raw trends graphs (Figure OC2) do not indicate a negative effect on unaffected workers. Nevertheless, to directly test for spillover effects, we define a dummy variable \( NearSubway_c \) which is equal to one if any of the other employees in the same department are affected by the opening of Line 15. Then, we estimate the specification in Equation (1) by adding this dummy and its interactions with the post dummy. We find no negative effect on coworkers or subordinates; instead, the results (Table OC15) suggest a large positive spillover effect (positive coefficient on the \( NearSubway_c * Post \)). We are cautious about interpreting these results, as only a small fraction (<10%) of workers do not have an affected colleague (mean of \( NearSubway_c \) is 0.91), so the variation available for identification appears to be limited. Nevertheless, these results suggest that the aggregate impact of the subway on worker performance is not attenuated due to negative spillover effects (and that baseline estimates may, in fact, be biased downward). Also, we verified that teams with more affected employees saw both a relative and absolute improvement in the team-level average total compensation after opening of the subway (Figure OC6).

To corroborate using stated company compensation policies, we followed up with human resources executives in both companies. Both noted that in order to foster cooperation and team spirit, they base employee bonus only on the absolute performance of the employees. They also
confirmed that they did not use fixed bonus pools for groups.

**E. Elasticity of Bonus to Commute Time**

Our setting allows for an estimate of the elasticity of bonus to commute time, which could be useful to assess benefits from other commute-saving interventions. One limitation is that we do not have a direct measure of commute time for the workers during the panel period, so we rely on imputing the pre-subway commute time for affected workers using measures of time saved by the subway. We examine a number of specifications, using different samples and alternative measures of commute time (Table OC16; the notes provide detailed variable and sample definitions). Overall, the results suggest a strong negative correlation between commute time and worker bonus pay as expected. We find the largest elasticity estimate of -0.33 (in Col 4 of Table OC16) when using self-reported commute time data; this implies that a 1% reduction in commute time increases bonus by 0.33%. The estimates are smaller (ranging from -0.078 to -0.163) in samples that include all workers from the baseline sample.

**4.4. Possible Channels for Observed Effects**

In this section, we explore evidence for these two plausible channels for the observed increase in performance: allocation of some of the commute time savings to work, and/or an increased intensity of effort (i.e., effort per unit of time) by the worker.

**A. Analysis of Attendance and Late/Early Arrivals**

To examine if the reduction in commute time translated into more time at the work place, we analyze two sources of data. First, we use daily data collected by the companies on whether the worker arrived late, left early, or took a leave of absence from work. We aggregate the
variables to the individual-month-level, and analyze a specification similar to Equation (1). In Cols 1 to 4 in Table 3, we find no significant differential effects on affected workers.

To check for the possibility of workers’ arriving early or staying late, we also collated data on employees swiping in and out office floors and in and out of the building. Unfortunately, our analysis of the data suggested significant measurement errors; e.g., a significant proportion of odd numbered swipes per day by employees, possibly because employees could enter or exit in groups. Nevertheless, assuming the nature of the measurement error stays the same across affected and unaffected employees (or over time), we expect a DID analysis to not be systematically biased. After cleaning the data, including eliminating single swipes and swipes very early in the morning or very late at night, and excluding weekends and holidays, we undertook two types of analyses.

First, we define attendance time as the time between first and last swipes, and then examine a regression of log monthly average attendance time per workday (Col 5 of Table 3), as well as the log time of arrival (Col 6) and the log time of leaving work (Col 7), using a DID specification similar to Equation (1). We find no significant effects of the opening of the subway. Col 8 confirms a significant increase in the log bonus for the subsample of workers for which we have cleaned attendance data.

We also examine the full kernel density distribution of time of arrival and time of leaving work (Figure 3). Consistent with the results in Table 3, we find no systematic evidence of a shift in distribution such that affected workers begin arriving earlier, or leaving later after the subway opens. In fact, the estimates suggest a slight shift right (later arrivals) for time of arrivals, and a slight shift left (earlier departure) for time of departure.
Taken together, our results suggest there was no increase in time spent in the office by affected workers.\textsuperscript{9} These results are consistent with a number of anecdotes from our interviews, where employees insisted that strong monitoring and work norms enforced very high compliance with expected daily arrival and exit time cut-offs, for all workers, before and after the opening of the subway. However, while we can rule out a role for extra time spent at the work place, our analysis does not rule out workers devoting some of the time saved on the commute to doing some work at home, or working during the better commute.

\textbf{B. Qualitative Evidence from Interviews}

Two rounds of detailed discussions and interviews of both non-marketing (first round) and marketing personnel (second round) yielded supplementary anecdotal evidence relating to potential mechanisms for some of our key findings, which we summarize briefly below.\textsuperscript{10}

i. Subway improves commute quality: The workers who were self-identified as taking the subway confirmed that the subway improves their commute. Interestingly, a main benefit they emphasized was that commute time became more regular and predictable. This concurs with research that finds predictability to be a key factor in commute-related stress (e.g., Evans, Wener, and Phillips, 2003). Also, the slight right shift of arrival time in Figure 3 is consistent a reduction in need to hedge unpredictable commutes. Some noted that their return commute was especially pleasant because the station was a terminus; this reduced

\textsuperscript{9} In the Online Appendix OA3, we present a model where effort is a product of work intensity and time at work. We show that the key predictions, including greater output and bonus, as well as a lower probability of exit, hold even if binding office hours are set by the firm, so that workers do not adjust the time at work.

\textsuperscript{10} Detailed notes from our two rounds of interviews are provided in the Online Appendix OF.
crowding and security clearance time relative to the next best subway option, which required an additional walk of about 1.1 kilometers, and had long security clearance delays.

ii. **Good commute provides non-pecuniary utility:** Through hypothetical questions posed to the interviewees, we confirmed that a superior commute is a valued job characteristic, with workers willing to give up a job that involved navigating bad traffic. Interviewees suggested the better commute improved their life satisfaction. This is also in line with research showing relation between shorter (e.g., Zhu et al., 2017) and rail (e.g., Meyer and Dauby, 2002) commutes with greater subjective wellbeing.\textsuperscript{11}

iii. **Workers allocate commute time savings to non-work tasks and rest:** Many of our affected interviewees indicated they did not spend extra time at work; instead, they used saved time for household related chores/errands, taking care of their parents/children, for exercise, relaxation or sleep.\textsuperscript{12} Marketing interviewees indicated that they were able to actually accomplish work during their improved commute on Subway Line 15. Interestingly even unaffected workers, when asked about hypothetical benefits from taking the subway,

\textsuperscript{11} Our employee survey asked about current work-life balance, and the workers’ recollection of their work-life balance levels in the period prior to the subway opening. Analysis of this data did not yield statistically significant effects for affected workers although estimates were positive for the full sample and non-managers (Appendix Table OB6). We believe this is likely because of significant limitations of recall data on happiness as documented by e.g., Ratner, Kahn, and Kahneman (1999). In addition, research shows that satisfaction levels revert to a stable set point over time (the “hedonic treadmill” effect) except for changes from major life events for some subsets of the population (e.g., Fujita and Diener 2005).

\textsuperscript{12} Recent research suggests a link between better sleep and labor productivity (Gibson and Shrader 2018), and between exercise and worker productivity (summarized in Lechner, 2015); thus better sleep and exercise provide potential pathways for our observed results.
reasoned that it would improve time management and reduce fatigue.

5. Analysis of Employee Exit, Hiring and Worker Sorting

5.1. Impact on Employee Exit Rate

We expect the opening of Subway Line 15 to lower probability of affected employees exiting the firms, because the improved commute increases both the non-pecuniary as well as wage benefits from the job. We investigate this prediction in this section. We define a dummy $Exit_t$ as equal to one in period $t$ if the employee is not in the company in month $t + 1$.

First, we examine some raw trends. In Figure 4, we find that the surviving fraction of unaffected workers was initially higher (in 2013) than that of affected workers (who had a longer commute before the subway opened), then was close to that of affected workers in the 12-month period prior to the opening of the subway. After the opening of the subway, the fraction of surviving unaffected workers gradually drops below that of affected workers, consistent with a relative decline in exit rate for affected workers.

We then examine the worker exit using duration and linear regressions models. In the first four columns of Table 4 we present results from proportional hazards models of the form

$$h_i(t) = h_0(t) \exp(\beta_1.NearSubway_i + \beta_2.T_t + \beta_3.NearSubway_i \times T_t + \gamma X_{it}).$$

where $NearSubway_i$ is a dummy for affected workers, and $T_t$ is a dummy for the period post subway opening, both as used in Section 4.1, and $X_{it}$ is a vector of controls including Age and Experience (time-varying), and Male dummy, Education, Number of Children, Manager dummy, Distance, Company 1 dummy, Married dummy, and New Employee dummy (all defined as at the start of the worker spell within the sample period). We define the duration $t$ as the amount of time from the start date of the worker (which is available for all workers);
the estimation allows for left censoring as we also specify the date that the worker is first observed in the sample panel period. In Cols 3 and 4, we add separate cubic polynomials in experience and age. In Cols 1 and 3 we use the semi-parametric Cox model, and in Cols 2 and 4 we check robustness to using a parametric (exponential) proportional hazards model.

The relative hazard for the Near Subway group was not significantly different (and in fact slightly higher) in the pre-subway period (positive $\beta_1$), but is significantly lower after subway opening (negative and significant $\beta_3$ coefficient) across all four specifications. The estimate from Col 2 implies a decline in hazard of exit of 49.8% ($1 - \exp(-0.69)$), while Col 4 implies a slightly smaller effect of 49.3% ($1 - \exp(-0.68)$).

In Cols 5 and 6 of Table 4, we present linear probability specifications, with a dummy for exit as the dependent variable. Both specifications include company-year-month fixed effects. Col 6 adds separate cubic polynomials in experience and age, as well as a fifth degree polynomial in tenure (Sandvik et al., 2021). These specifications yield similar estimates: we find about 1.1% lower probability of exit per month for affected workers in the post-subway opening period. Additional analysis showed larger and statistically stronger reduction in exit propensity for the group with above median commute time savings (Appendix Table OD1).

Given that the opening of the subway could be anticipated, and that changes in exit propensity may be slow to manifest as affected workers experience and learn the benefits of

---

13 The coefficients on controls were in line with expectations: exit hazard is consistently higher for males, lower for individuals with more children, and higher for new employees. Interestingly, exit hazard is strongly positively correlated with distance, consistent with the baseline finding that reduction of commute reduces exit hazard.
the subway, the results here should be interpreted with caution. In particular, while tests show a significant structural break in the differential exit rate for affected workers one month after the opening of the subway, the results for the most significant break are consistent with worker retention effects manifesting more strongly after the affected workers have experienced the commute benefits for a few months (Appendix Table OD2).

5.2. Impact on Hiring and Worker Sorting

We examine changes in worker quality, and worker sorting into and out of the firm, and towards and away from Subway access (for brevity, results are in the online appendix).

A. Improvement in Quality of Hires

We examine whether the companies are able to attract better-performing employees from locations close to the subway because of the non-pecuniary utility from the easy commute, for the same salary offers as made to unaffected applicants. Specifically, for the sample of employees hired after the opening of the subway, we regress performance scores (available for 2015 and 2016, and used in robustness checks above) on NearSubway after controlling for different forms of base salary and year-month fixed effects. As expected, we find that conditional on total compensation, affected employees hired after the subway opening have, on average, significantly higher performance scores (Cols 1-3, in Appendix Table OB7, Panel A). While not an ideal placebo check (as some selection could also happen before the anticipated subway opening), we find that the coefficients of NearSubway are smaller and statistically insignificant (Cols 4-6, in Appendix Table OB7, Panel A) for workers from affected locations hired before the subway opening. These results, as well as other analysis (discussed in Appendix OB4), suggest that the companies are able to hire higher-quality workers from the
affected applicant pool after the subway opened.

**B. Additional Analysis of Worker Sorting**

In this section we explore how the composition of workers at the firm changed following the subway opening. We examine four trend graphs (in Appendix Figure OB2), which show that: (a) the fraction of new hires that have commute access to the Subway are not systematically higher after the opening of the subway.\(^{14}\) (b) The fraction of leavers who have commute access to the subway is indeed lower after the subway opens; (c) Home moves by workers appear to be influenced by factors other than the access to the subway (e.g., access to schools, mentioned by our interviewees), as the number of movers to subway access is small relative to the total (which is itself modest, less than four movers in most quarters). (d) The lack of increase in fraction of hires with subway access, combined with moves away from the subway, appear to offset the positive sorting from exits, so that the overall fraction of workers affected by the subway stays relatively flat, except towards the end of the sample period.

**6. External Validity: Check Using Broader Set of Subway Construction Event Studies**

We investigate whether access to a new, nearby subway increases firm market value. In addition to being of independent interest, this analysis for a large sample of firms could also be informative about the broader validity of our earlier results, assuming that an important driver for firm value is improved worker performance and retention.

\(^{14}\) Possible explanations for this result are that (i) the firms do not fully realize the differential benefit from hiring affected workers, or (ii) do not emphasize this in recruiting (as acknowledged by an HR executive), or (iii) the number of hires is low, and availability of potential affected hires is limited.
We undertake supplementary data collation on 16 subway construction commencement events and a related large sample of affected and unaffected companies. The Online Appendix includes a detailed description of this analysis (Appendix OC14) and a table with the full list of companies impacted by each of the events (Table OC17). We find evidence that initiation of construction of a new subway station near the company increases firm value, as reflected in positive abnormal returns around the event date (Table OC18). Combined with our earlier findings, these results suggest that both workers and firms benefit from access to the subway.

7. Aggregate Surplus to Workers and the Firms from Subway Opening

While our data is best suited to understanding whether and how much an improvement in commute affects worker performance and retention, here we cautiously explore impacts on the total surplus to the worker and to the firm.

A. Surplus to Workers and Firms, and Rent-sharing Elasticity

Worker Surplus Including Imputed Disutility of Commute: To examine the total worker surplus inclusive of the value of commute time savings, we define worker welfare or commute-adjusted-compensation as total compensation of worker \( i \) in month \( t \) less commute cost, where commute cost is equated to half of the wage rate following the norm in the literature per Redding and Turner’s handbook chapter (2015, pp 1353). We find that the new subway improved affected worker welfare by about 5.6% (Col 1 of Table OE1).

Gains to the Companies based on Rent-Sharing Elasticity Estimates in the Literature: Based on the wage elasticities to value added of 0.05-0.15 reported in a paper by Card et al. (2018) that reviews the rich literature on rent sharing, the gains to the firms in log value added per affected worker would be 20 to 6.7 times the log change of 2.9% in total worker compensation.
Estimated Rent-Sharing Elasticities Based Marketing Workers' Data: We attempt to contribute to the rent sharing literature using data available for marketing employees (see Online Appendix Tables OE3 and OE4 for details). Combining the estimated effect on individual-level log sales (0.79, Appendix Table OC2) with the estimate of the effect of subway opening on log bonus of marketing workers (0.094, Appendix Table OC14, Col 5) implies a bonus elasticity of sales of about 0.12 (0.094/0.79). The point estimate of 0.027 for total monetary compensation for marketing workers (Appendix Table OE2, Col 1) implies a total compensation elasticity to sales of 0.034, and the log welfare measure change of 0.051 (Appendix Table OE2, Col 2) implies a worker welfare elasticity to sales estimate of 0.065. Our estimates of elasticity of bonus and total welfare to sales are reassuringly within the 0.05-0.15 range of preferred wage elasticities reported by Card et al. (2018).

B. A Back-of-the-Envelope Cost-Benefit Assessment

Finally, in Table 5, we undertake a limited back-of-the-envelope cost-benefit comparison, relating the estimate of the gains to workers in the neighborhood of the subway, to available estimates of the cost of construction of the new subway station. Using data from the National Tax Survey, we estimate that the total annual income of workers in firms located within a 500-meter radius of the new Subway station was about RMB10 billion; then using an estimated impact on total income of 2.9% (from Col 5 of Table 2), and affected proportion of workers of 39% (from Col 2, Row 1 of Table 1), and assuming the benefits decline to zero at the radius of 500m, yields an estimated annual income gain of RMB56.6 million. Further, assuming commute time saved also declines from a maximum of 25.3 minutes at our firms (twice median
commute time savings from Col 3, Row 2 of Table 1) to zero at the radius of 500m, and valuing commute time at half of the compensation rate (per Redding and Turner 2015), we obtain annual commute time savings of RMB51.5 million. Using data from Wikipedia and other sources (listed in Col 3 of Table 5), we yield a total construction cost of about RMB1.4 billion for the new subway station and the additional line to the next nearest subway station. This implies an annual return in income gains to workers of about 4% per year (56.6/1417), and benefits in commute saving of about 3.6% (51.5/1417), for a combined annual benefit of 7.6% of construction costs.\textsuperscript{15} As a benchmark related to borrowing cost for the Chinese government, the prevailing yield on 10-year Chinese government bonds was in the range of 3.28% to 5% between 2011 and 2014 (per data from the Wall Street Journal: Market’s site).

Thus, our rough estimates suggest the annual benefits to workers exceed the annualized costs of government borrowing that could be used to finance the subway construction. However, a number of factors suggest strong caution in interpreting these results. First, what we undertake is a marginal cost-benefit analysis which compares additional benefits from the subway station in a dense and high-income downtown neighborhood to the marginal cost of construction of the new subway station; this ignores the costs of construction of the remaining portion of Subway Line 15 and the various connected lines that help the affected workers on their commute. Second, the estimate of income effects in Table 2 is noisy, and hence the mean

\textsuperscript{15} An alternative approach to estimating the annual income of affected workers based on the number of commuters using the rail during peak hours (discussed in the notes to Table 5) yields a similar estimate. Our estimation assumes that annual operating costs are covered by ticket costs (and that these are roughly equal to the cost of alternatives for commuters).
estimate of income gains should be viewed with caution. Third, the magnitude of benefits from a percentage increase in income, as well as from commute time savings, depends crucially on the annual income of the affected workers. Workers in our sample report average annual income of RMB140,000 over our panel period, while average per capita GDP in China was only around RMB50,000 in 2015, per World Bank data. Finally, the time savings for other locations with less congested road traffic could be significantly lower than in our context.

8. Conclusion

Using “insider” company administrative and survey data, we explore how employees responded to the availability of improved public commute, through the opening of a subway station close to their workplace. We posit that a better commute enhances worker effort (through improved wellbeing and/or from freeing up commute time) and non-pecuniary utility from the job. The subway opening is therefore expected to increase worker bonus pay, and reduce propensity to exit the firm (and enhance quality of hires from subway-accessible locations). We find that these expectations are borne out in the data, for affected workers (defined as those for whom the new Subway Line 15 provides the fastest public transport commute route from home), relative to others.

A careful exploration of data on attendance reveals that there was no allocation of time saved from commute towards time at the workplace. This suggests that a main pathway for the observed improvement in bonus and increased retention is through an improvement in wellbeing, which in turn prompted greater worker effort (Porto, 2016). A rough back-of-the-envelope estimate suggests that the total annual benefits to the workers affected by the new subway were about 7.6% of construction costs.
Our study is limited in its focus on the relative performance, retention, and hiring effects for affected workers. We do not examine other benefits from the subway line including potential pass-through of productivity gains to consumers, and to other input suppliers (e.g., landlords). There are also potential benefits from reduction in air pollution from reduced road traffic congestion (Anderson, 2014; Chen and Whalley, 2012), and long-run gains from entry of new firms and workers facilitated by new subway lines, which we do not address.

**References**


Figure 1: Subway Line 15 and Connected Lines
Figure 2: Trends in Differentials for Log (Bonus$^T$)

Notes: Each point is the coefficient on a quarter dummy interacted with NearSubway, which captures the difference in the Log(Bonus$^T$) in the specific year-quarter between the affected workers compared to the control group. The reference quarter is the fourth quarter of 2014 (hence normalized to zero). The error bar shows the 95% confidence interval, based on two-way (at individual and year-month level) clustered standard errors. The vertical dashed line indicates the opening of Subway Line 15.
Figure 3: Time of Arrival and Leaving

Panel A. Time of Arrival

(i) Affected Workers
(ii) Unaffected Workers

Panel B. Time of Leaving

(i) Affected Workers
(ii) Unaffected Workers

Notes: This figure shows the density of the time of arrival and leaving for the two periods (before and after the subway line 15 opening). We collect the daily swiping-in and swiping-out data for the two companies from Jan. 1, 2013 to Dec. 31, 2016. The vertical line represents the official work-start time in Panel A and the official work-end time in Panel B.
Figure 4: Survival Rate for Initial Sample of Workers

Notes: This figure presents the fraction of workers surviving from the sample of workers present in the companies at start of our sample period (Jan. 2013). Thus, the survival rate is 100% in January 2013 by construction.
<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>Mean</th>
<th>Mean</th>
<th>Diff</th>
<th>p</th>
<th>Mean</th>
<th>Mean</th>
<th>Diff</th>
<th>p</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Media</td>
<td>Media</td>
<td>Std.</td>
<td></td>
<td>Media</td>
<td>Media</td>
<td>Std.</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Obs.</strong></td>
<td>14,807</td>
<td>14,807</td>
<td>14,807</td>
<td>14,807</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>14,807</td>
<td>14,807</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>NearSubway Dummy</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.39</td>
<td>0.00</td>
<td>0.49</td>
<td></td>
<td>1.00</td>
<td>0.00</td>
<td>1.00</td>
<td>0.00</td>
<td>NA</td>
</tr>
<tr>
<td><strong>TimeSaved (one-way,</strong></td>
<td>5,689</td>
<td>11,576</td>
<td>11,576</td>
<td>11,576</td>
<td>21.77</td>
<td>12.67</td>
<td>24.88</td>
<td></td>
<td>20.91</td>
<td>0.00</td>
<td>20.91</td>
<td>0.00</td>
<td>NA</td>
</tr>
<tr>
<td><strong>Distance (kilometer)</strong></td>
<td>14,807</td>
<td>14,807</td>
<td>14,807</td>
<td>14,807</td>
<td>10.50</td>
<td>8.65</td>
<td>8.73</td>
<td></td>
<td>11.99</td>
<td>10.20</td>
<td>1.78</td>
<td>0.06</td>
<td>80.16</td>
</tr>
<tr>
<td><strong>Bonus (RMB)</strong></td>
<td>14,807</td>
<td>14,807</td>
<td>14,807</td>
<td>14,807</td>
<td>4,945</td>
<td>3,305</td>
<td>5,526</td>
<td></td>
<td>4,735</td>
<td>4,744</td>
<td>-8.66</td>
<td>0.99</td>
<td>NA</td>
</tr>
<tr>
<td><strong>Log (Bonus)</strong></td>
<td>14,807</td>
<td>14,807</td>
<td>14,807</td>
<td>14,807</td>
<td>9.30</td>
<td>9.23</td>
<td>0.39</td>
<td></td>
<td>9.24</td>
<td>9.25</td>
<td>-0.01</td>
<td>0.81</td>
<td>NA</td>
</tr>
<tr>
<td><strong>Log(Total Income Net of</strong></td>
<td>14,807</td>
<td>14,807</td>
<td>14,807</td>
<td>14,807</td>
<td>8.55</td>
<td>8.61</td>
<td>0.96</td>
<td></td>
<td>7.88</td>
<td>8.02</td>
<td>-0.15</td>
<td>0.37</td>
<td>NA</td>
</tr>
<tr>
<td><strong>Log (Total Income)</strong></td>
<td>14,807</td>
<td>14,807</td>
<td>14,807</td>
<td>14,807</td>
<td>9.77</td>
<td>9.70</td>
<td>0.35</td>
<td></td>
<td>9.58</td>
<td>9.62</td>
<td>-0.03</td>
<td>0.47</td>
<td>NA</td>
</tr>
</tbody>
</table>

Notes: NearSubway, TimeSaved and Distance, are defined using the Baidu API (see online Data Appendix). The propensity model specification is presented in Online Appendix Table OB3. Bonus and income variables are deflated by CPI (2013 base year). To retain information on negative and zero bonus income in logs, the Bonus (and for consistency Total Income) are a simple linear transformation of bonus and total income by adding a constant (6905.22) such that the minimum value of the adjusted bonus is equal to one.
Table 2: Impact of the Opening of Subway Line 15 on Employee Compensation

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
<th>(7)</th>
<th>(8)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Log (Total Income)</td>
<td>Log(Total Income Net of Bonus)</td>
<td>Bonus</td>
<td>Log (Bonus)</td>
<td>Log (Total Income)</td>
<td>Log(Total Income Net of Bonus)</td>
<td>Bonus</td>
<td>Log (Bonus)</td>
</tr>
<tr>
<td></td>
<td>(Total Income)</td>
<td>Net of Bonus</td>
<td>(Total Income)</td>
<td>Net of Bonus</td>
<td>(Total Income)</td>
<td>Net of Bonus</td>
<td>(Total Income)</td>
<td>Net of Bonus</td>
</tr>
<tr>
<td>Panel A: Baseline sample</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Self-reported NearSubway × Post</td>
<td>0.027</td>
<td>-0.062</td>
<td>957**</td>
<td>0.067*</td>
<td>(0.021)</td>
<td>(0.071)</td>
<td>(475)</td>
<td>(0.033)</td>
</tr>
<tr>
<td>NearSubway × Post</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.029*</td>
<td>-0.005</td>
<td>636*</td>
<td>0.049*</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(0.016)</td>
<td>(0.055)</td>
<td>(370)</td>
<td>(0.027)</td>
</tr>
<tr>
<td>Adj. R-Squared</td>
<td>0.85</td>
<td>0.59</td>
<td>0.66</td>
<td>0.72</td>
<td>0.85</td>
<td>0.56</td>
<td>0.67</td>
<td>0.72</td>
</tr>
<tr>
<td>Obs</td>
<td>9,529</td>
<td>9,529</td>
<td>9,529</td>
<td>9,529</td>
<td>14,807</td>
<td>14,807</td>
<td>14,807</td>
<td>14,807</td>
</tr>
<tr>
<td>N (affected group)</td>
<td>62</td>
<td>62</td>
<td>62</td>
<td>62</td>
<td>144</td>
<td>144</td>
<td>144</td>
<td>144</td>
</tr>
<tr>
<td>N (unaffected group)</td>
<td>160</td>
<td>160</td>
<td>160</td>
<td>160</td>
<td>233</td>
<td>233</td>
<td>233</td>
<td>233</td>
</tr>
<tr>
<td>Change on level variable</td>
<td>511</td>
<td>-431</td>
<td>957</td>
<td>803</td>
<td>542</td>
<td>-34</td>
<td>636</td>
<td>581</td>
</tr>
<tr>
<td>Effect as % of SD of Dep Var</td>
<td>7.9%</td>
<td>-6.3%</td>
<td>17.3%</td>
<td>17.0%</td>
<td>8.2%</td>
<td>-0.5%</td>
<td>11.5%</td>
<td>12.6%</td>
</tr>
<tr>
<td>Panel B: Propensity-score matched sample</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Self-reported NearSubway × Post</td>
<td>0.031</td>
<td>-0.073</td>
<td>1,051**</td>
<td>0.074**</td>
<td>(0.023)</td>
<td>(0.079)</td>
<td>(504)</td>
<td>(0.036)</td>
</tr>
</tbody>
</table>
 NearSubway × Post | 0.027 | -0.020 | 694* | 0.050*  
(0.018) | (0.063) | (403) | (0.029)  
Adj. R-Squared | 0.86 | 0.61 | 0.7 | 0.73 | 0.85 | 0.58 | 0.69 | 0.72  
Obs | 7,237 | 7,237 | 7,237 | 7,237 | 11,086 | 11,086 | 11,086 | 11,086  
N (affected group) | 59 | 59 | 59 | 59 | 141 | 141 | 141 | 141  
N (control group) | 109 | 109 | 109 | 109 | 141 | 141 | 141 | 141  
Change on level variable | 584 | -513 | 1051 | 875 | 496 | -136 | 694 | 579  
Effect as % of SD of Dep Var | 8.8% | -7.3% | 19.9% | 19.2% | 7.5% | -2.2% | 13.2% | 13.3%  

Note: See Section 3.3 for definitions of the dependent variables. Self-reported NearSubway is a dummy defined as 1 for those workers reporting that they use the Subway Line 15, per responses to our survey of workers. Accordingly, the sample includes only existing workers in June 2017 who responded to our survey for those specifications. NearSubway is a dummy variable equal to one if the fastest public transport route from the worker’s home address to office contains Subway Line 15, otherwise zero. In Panel B, the sample consists of affected workers and their nearest propensity-score matched neighbor, per procedure in Ho et al. (2011). All specifications include Individual and Company-Year-Month fixed effects. Two-way (at individual and year-month level) clustered standard errors are reported in parentheses. *, **, and *** denote statistical significance at 10%, 5%, and 1%, respectively.
Table 3: Effect of the Subway on Employees' Attendance and Time at Work

<table>
<thead>
<tr>
<th></th>
<th>(1) Log(Late for Work)</th>
<th>(2) Log(Leave Early)</th>
<th>(3) Log(Sick Leave)</th>
<th>(4) Log (Personal Leave)</th>
<th>(5) Log(Attendance Time per Workday)</th>
<th>(6) Log(Time of Arrival)</th>
<th>(7) Log(Time of Leaving)</th>
<th>(8) Log(Bonus) for attendance sample</th>
</tr>
</thead>
<tbody>
<tr>
<td>NearSubway × Post</td>
<td>-0.001</td>
<td>0.003</td>
<td>0.000</td>
<td>0.005</td>
<td>0.104</td>
<td>-0.006</td>
<td>0.007</td>
<td>0.083***</td>
</tr>
<tr>
<td></td>
<td>(0.018)</td>
<td>(0.003)</td>
<td>(0.009)</td>
<td>(0.017)</td>
<td>(0.101)</td>
<td>(0.009)</td>
<td>(0.009)</td>
<td>(0.030)</td>
</tr>
<tr>
<td>Adj. R-Squared</td>
<td>0.275</td>
<td>0.047</td>
<td>0.071</td>
<td>0.103</td>
<td>0.672</td>
<td>0.557</td>
<td>0.527</td>
<td>0.735</td>
</tr>
<tr>
<td>Obs</td>
<td>14,742</td>
<td>14,742</td>
<td>14,742</td>
<td>14,741</td>
<td>4,486</td>
<td>4,486</td>
<td>4,486</td>
<td>11,534</td>
</tr>
<tr>
<td>N (affected group)</td>
<td>144</td>
<td>144</td>
<td>144</td>
<td>144</td>
<td>115</td>
<td>115</td>
<td>115</td>
<td>115</td>
</tr>
<tr>
<td>N (control group)</td>
<td>233</td>
<td>233</td>
<td>233</td>
<td>233</td>
<td>164</td>
<td>164</td>
<td>164</td>
<td>164</td>
</tr>
</tbody>
</table>

Notes: Each dependent variable is in terms of the number of days of the month. For Cols 5-7, we collect the daily swiping-in and swiping-out data for the two companies, and calculate the attendance time as the difference between the time of the first swiping-in and the last swiping-out.

Attendance time per workday is defined as the total monthly attendance time in the workdays divided by the number of workdays in the month.

Col 8 uses the same sample of employees in Cols 5-7 but covers more months. All specifications include Individual and Company-Year-Month fixed effects. Two-way (at individual and year-month level) clustered standard errors are reported in parentheses. *, **, and *** denote statistical significance at 10%, 5%, and 1%, respectively.
### Table 4: Effects of the Opening of Subway Line 15 on Employee Exit

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Hazard Models</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Cox</td>
<td>Exponential</td>
<td>Cox</td>
<td>Exponential</td>
<td>Exit Dummy (OLS)</td>
<td></td>
</tr>
<tr>
<td>NearSubway × Post</td>
<td>-0.66***</td>
<td>-0.69***</td>
<td>-0.65***</td>
<td>-0.68***</td>
<td>-0.011***</td>
<td>-0.011**</td>
</tr>
<tr>
<td></td>
<td>(0.23)</td>
<td>(0.23)</td>
<td>(0.22)</td>
<td>(0.22)</td>
<td>(0.004)</td>
<td>(0.004)</td>
</tr>
<tr>
<td>NearSubway</td>
<td>0.14</td>
<td>0.11</td>
<td>0.18</td>
<td>0.14</td>
<td>0.001</td>
<td>0.001</td>
</tr>
<tr>
<td></td>
<td>(0.15)</td>
<td>(0.15)</td>
<td>(0.15)</td>
<td>(0.15)</td>
<td>(0.003)</td>
<td>(0.003)</td>
</tr>
<tr>
<td>Post</td>
<td>0.37**</td>
<td>0.31**</td>
<td>0.39***</td>
<td>0.31**</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.15)</td>
<td>(0.14)</td>
<td>(0.15)</td>
<td>(0.14)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Controls</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Additional Controls</td>
<td>No</td>
<td>No</td>
<td>Cubic in experience, Cubic in age</td>
<td>No</td>
<td>Cubic in experience, Cubic in age, 5th degree polynomial in tenure</td>
<td></td>
</tr>
<tr>
<td>Individual fixed effects</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Company-Year-Month fixed effects</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Observations</td>
<td>17,718</td>
<td>17,718</td>
<td>17,718</td>
<td>17,718</td>
<td>17,718</td>
<td>17,718</td>
</tr>
<tr>
<td>Log Likelihood</td>
<td>-1540</td>
<td>-489.6</td>
<td>-1530</td>
<td>-483.8</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Notes: The first four Cols represent proportional hazards models, with the workers at risk for exit from their start date in the companies, and in the data from start of the panel period. Cols 1 and 3 use a Cox Proportional Hazards specification, while Cols 2 and 4 present an exponential model. Controls include dynamic variables (Age and Experience) and static (defined at start of panel) controls (Education, Number of Children, Male Dummy, Manager Dummy, Distance, Company 1 Dummy, Married Dummy, and New Employee Dummy). Robust standard errors (clustered at the individual level) are reported in parentheses. For Cols 5 and 6, the dependent variable Exit is an indicator which equals 1 if the employee exited the company, and zero otherwise; Employee exit is recorded in month t if month t is the last month for this employee in the company; both columns use OLS with Company-Year-Month fixed effects.
Table 5: Back-of-the Envelope Cost-Benefit Calculations

<table>
<thead>
<tr>
<th>Description</th>
<th>Estimate</th>
<th>Source/Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Estimate of aggregate benefit to workers in subway neighborhood:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Estimated effect on total income of affected worker (A)</td>
<td>0.029</td>
<td>From Col 5 of Table 2</td>
</tr>
<tr>
<td>Fraction of workers who are affected (B)</td>
<td>0.39</td>
<td>From Col 2, Row 1 of Table 1</td>
</tr>
<tr>
<td>Radius of benefit from the subway (in meters)</td>
<td>500</td>
<td>Half of distance to the next closest subway station (Line 13)</td>
</tr>
<tr>
<td>Annual income of workers within 500m of subway (C), in million RMB</td>
<td>10,000</td>
<td>Authors’ calculations from tax survey data, aggregated using survey sampling weight factor of 10 (Liu and Mao, 2019)</td>
</tr>
<tr>
<td>Linear adjustment factor (D)</td>
<td>0.5</td>
<td>Linear decline in benefits/time savings, going to zero at benefit radius</td>
</tr>
<tr>
<td><strong>Annual income gains (E=A × B × C × D), in million RMB</strong></td>
<td>56.6</td>
<td></td>
</tr>
<tr>
<td>Time saved per day (F) (in minutes) for workers next to Subway</td>
<td>25.34</td>
<td>2×median one-way time saved, from Col 3, Row 2 of Table 1</td>
</tr>
<tr>
<td>Working time in minutes per day (G)</td>
<td>480</td>
<td>Assuming 8 hour working day</td>
</tr>
<tr>
<td>Adjustment factor for value of commute time (H)</td>
<td>0.5</td>
<td>Valued at 50% of wage rate, per Redding and Turner (2015, pp 1353)</td>
</tr>
<tr>
<td><strong>Annual commute savings (I = F × B × C × D × H/G), in mn RMB</strong></td>
<td>51.5</td>
<td>Income per working minute of affected workers (B×C/(Number of working days×G)) × average commute time saved by affected workers (F×D×Number of working days) adjusted by commute value factor (H)</td>
</tr>
<tr>
<td><strong>Total annual benefits (J= E+I), in million RMB</strong></td>
<td>108.1</td>
<td></td>
</tr>
<tr>
<td>Estimate of total construction costs for new subway station</td>
<td></td>
<td></td>
</tr>
<tr>
<td>----------------------------------------------------------</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Construction cost of the subway station (K), in million RMB</td>
<td>273</td>
<td>Internal document of China Railway Construction Group Co., LTD.</td>
</tr>
<tr>
<td>Distance to next subway station (L), in km</td>
<td>1.144</td>
<td>Wikipedia (2021a), verified using Baidu.</td>
</tr>
<tr>
<td>Construction cost per km of subway line (M), in million RMB</td>
<td>1000</td>
<td>National Development and Reform Commission (NDRC) cited in ECNS (2015); Wikipedia (2021a), and WSJ (2014).</td>
</tr>
<tr>
<td><strong>Total construction costs (N = K + L × M), in million RMB</strong></td>
<td>1417</td>
<td></td>
</tr>
</tbody>
</table>

| Annual return in terms of income gains (E/N) | 3.99% |
| Annual return in terms of commute savings (I/N) | 3.63% |
| **Annual return in terms of total benefits (J/N)** | 7.63% |

| China 10-year government bond yields | 3.28% to 5% | Data for 2011 to 2014, from Wall Street Journal: Markets (accessed July 7, 2021); verified using Barrons.com |

Note: The total annual income of affected workers has a substantial impact on the benefit calculations. Relying on the tax survey data, the estimate in this table is 3,890 million RMB (B × C). An alternative estimate using train capacity and frequency yields a similar figure.

Specifically, 19 trains arrive at Tongfang Plaza station during peak hours (7:30am to 9:30am, Source: bjsubway.com) × rated capacity of 1428 passengers per train (Source: Wikipedia 2021b) × Annual Salary of RMB 140,000 (from Table 1) = RMB3,798 million. While actual passenger numbers may exceed the rated capacity at the peak of rush hour, it may be lower at non-peak times.