USC FBE APPLIED ECONOMICS WORKSHOP
presented by: John Roberts
Friday, Mar. 14, 2014
10:30 am - 12:00 pm; Room: **JKP-210**

DOES WORKING FROM HOME WORK?

EVIDENCE FROM A CHINESE EXPERIMENT

Nicholas Bloom[a], James Liang[b], John Roberts[c] and Zhichun Jenny Ying[d]

February 24, 2014

Abstract:
About 10% of US employees now regularly work from home (WFH), but there are concerns this can lead to "shirking from home." We report the results of a WFH experiment at CTrip, a 16,000-employee, NASDAQ-listed Chinese travel agency. Call center employees who volunteered to WFH were randomly assigned to work from home or in the office for 9 months. Home working led to a 13% performance increase, of which about 9% was from working more minutes per shift (fewer breaks and sick-days) and 4% from more calls per minute (attributed to a quieter working environment). Home workers also reported improved work satisfaction and experienced less turnover, but their promotion rate conditional on performance fell. Due to the success of the experiment, CTrip rolled out the option to WFH to the whole firm and allowed the experimental employees to re-select between the home or office. Interestingly, over half of them switched, which led to the gains from WFH almost doubling to 22%. This highlights the benefits of learning and selection effects when adopting modern management practices like WFH.

Keywords: working from home, organization, productivity, field experiment, and China

Conflict of interest statement: No funding was received from CTrip. James Liang is the co-founder of Ctrip. During the experiment we report here he was non-executive Chairman of Ctrip. Since the end of the experiment he has returned to Ctrip as CEO. No other co-author has any financial relationship (or received any funding) from CTrip. The results or paper were not pre-screened by anyone.

[a] Stanford Economics, SIEPR, CEPR and NBER; [b] Beijing University Guanghua Management School and CTrip, [c] Stanford GSB and SIEPR, [d] Stanford Economics and Cornerstone Research.

I. INTRODUCTION

Working from home (also called telecommuting or telework, but hereafter referred to as "WFH") is becoming an increasingly common practice. In the United States, the proportion of employees that primarily works from home has more than tripled over the past 30 years, from 0.75% in 1980 to 2.4% in 2010 (Census, 2010).[1] At the same time, the wage discount (after controlling for observables) from working exclusively at home has fallen, from 30% in 1980 to zero in 2000 as WFH moved from being predominantly found in low-skilled jobs to encompassing a wider set of occupations (Oettinger, 2010). Home-based workers now span a wide spectrum of jobs, ranging from sales assistants to managers and software engineers, with a correspondingly wide range of incomes (Figure 1).

Having employees work from home raises two major issues. First, is it a useful management practice? This is an important question that lacks systematic evidence or consensus.[2] Thus, even within a single industry, practices often vary dramatically. For example, jetBlue Airlines' call-center employees all work from home, American Airlines does not allow any home work, and United Airlines has a mix of practices. More generally, Bloom et al. (2009) reported that 30% of U.S. and 33% of European manufacturing firms offer opportunities for at least some managers to work from home, with wide variation in adoption rates within every 3-digit SIC code surveyed.

The second issue relates to the concerns over deteriorating work-life balance in the US and other developed economies and to the potential of working from home to help address this. The number of households in the US with both parents working has increased from 25% in 1968 to 48% in 2008 (Council of Economic Advisors, 2010). The increasing pressure for both parents to work is leading governments in the US and Europe to investigate ways to promote work-life balance, with again a shortage of evidence:

> *A factor hindering a deeper understanding of the benefits and costs of flexibility is a lack of data on the prevalence of workplace flexibility and arrangements, and more research is needed on the mechanisms through which flexibility influences workers' job satisfaction and firms' profits to help policy makers and managers alike.*
>
> (Council of Economic Advisors, 2010, Executive Summary)

The efficacy of WFH as a management practice was what concerned CTrip, China's largest travel agency, with 16,000 employees and a NASDAQ listing. Its senior management were interested in allowing its Shanghai call-center employees to work from home because they perceived potential benefits from reducing office rental costs, which were increasing rapidly due to the booming real estate market in Shanghai. They also thought that allowing WFH might reduce the high attrition rates the firm was experiencing by saving the employees from long commutes. But the managers worried that allowing employees to work at home, away from the direct oversight of their supervisors, would lead to a large increase in shirking. The call center workforce was mainly younger employees, many of whom might well have struggled to remain focused working from home without direct supervision.

---

[1] This excludes self-employed workers. In the E.U. in 2005, 7.0% of employees worked "*with a computer away from the employer's premises at least a quarter of the time*", and 1.7% did so "*almost all the time*" (EIROnline, 2010).
[2] See, for example, Bailey and Kurland (2002).

Given the uncertainty surrounding the effects of working from home in the research literature as well as in practice, the firm's senior management decided to run a randomized controlled trial. The authors assisted in designing the experiment and had complete access to the resulting data, as well as to data from surveys conducted by the firm. We also conducted various surveys ourselves and numerous interviews with employees, line supervisors and senior management.

In summary, CTrip decided to run a nine-month experiment on working from home. They asked the 996 employees in the airfare and hotel departments of the Shanghai call center whether they would be interested in working from home four days a week, with the fifth day in the office as usual. Approximately half of the employees (503) were interested, particularly those who were married, had children and faced long commutes to work. Of these, 249 were qualified to take part in the experiment by virtue of having at least six months tenure, broadband access and a private room at home in which they could work. After a lottery draw, those of these 249 with even-numbered birthdays were selected to work from home, while those with odd-numbered birthdates stayed in the office to act as the control group.

Office and home workers used the same IT equipment, faced the same work order flow from a common central server, carried out the same tasks and were compensated under the same pay system, which included an element of individual performance pay. Hence, the only difference between the two groups was the location of work.[3] This allows us to isolate the impact of working from home versus other practices that are often bundled alongside this practice in attempts to improve work-life balance, such as flexible work hours.

We found several striking results. First, the performance of the home workers went up dramatically, increasing by 13% over the nine months of the experiment. This improvement came mainly from a 9% increase in the number of minutes they worked during their shifts (i.e., the time they were logged in to take calls). This was due to a reduction in breaks, time off and sick-days taken by the home workers. The remaining 4% improvement came from home workers increasing the number of calls per minute worked. In interviews, the workers attributed this gain to the quieter working conditions at home. Second, there appear to be no spillovers to the rest of the group. Comparing the control group to similar workers in CTrip's other call center in Nan Tong we see no performance drop despite the control group's having lost the treatment lottery. Third, attrition fell sharply among the home workers, dropping by 50% versus the control group. Home workers also reported substantially higher work satisfaction and had more positive attitudinal survey outcomes. Fourth, one downside of WFH appears to be that, conditional on performance, it reduced rates of promotion by about 50%.

There are some obvious concerns with these results. First, was quality sacrificed for quantity by the home workers? Using two different quality metrics we found no impact on quality of home working. Second, could the results be driven by the control workers' becoming frustrated by losing the randomization lottery and then performing worse? Fortunately, CTrip had a second large call center in Nan Tong that was not involved in the experiment. We compared the Shanghai-based

---

[3] This of course had implications that were potentially relevant to the experiment. In particular, employees at home did not have on-going, immediate contact with their supervisors and they worked in a different environment than those in the office. We discuss these points more below.

control group to similar employees in Nan Tong. We found almost identical results. Similarly we could compare the control group workers to eligible *non-volunteers* – employees who met the eligibility requirements but never wanted to work from home – and we again found almost identical results. Third, perhaps our results are driven by attrition bias. It turns out that in fact our results probably are biased by attrition, but biased *downwards*, so the true impact of WFH is probably substantially larger. In both control and treatment groups worse performing employees tended to quit more rapidly, but the since the quit rates were twice as large in the control group this generated a larger upwards bias on the performance measures for the control group. Using the approach of Lee (2008), we found an upper bound on the treatment effects that is about 50% higher than our baseline 13% estimate.

Finally, at the end of the experiment, the firm estimated it would save about $2,000 per year per employee working at home, leading it to offer the option to work from home to the entire firm. It also allowed the treatment and control groups to re-select their working arrangements. Surprisingly, over half of all the employees changed their minds, indicating the extent of employee learning about their own suitability for working from home. In particular, two thirds of the control group (who initially had all volunteered to work from home 10 months earlier) decided to stay in the office, citing concerns over the loneliness of home working. In reverse, half of the treatment group changed their minds and returned to the office – on average, those who had performed relatively badly at home, but also ones who found the lack of social contact particularly costly.

This learning and re-selection led to the impact from working at home rising to 22%, almost double the direct experiment effect of 13%. The reason was strong selection effects: workers with worse performance at home over the 9 month experiment period returned to the office, while those who performed well at home stayed at home. Strikingly, this ratio of selection plus direct effects (22%) to direct effects (13%) is similar to the 2:1 ratio in Lazear's well-known study of introducing piece-rate pay in Safelite Auto Glass (Lazear 2000). This highlights how selection effects of employees across different working practices are an important part of the impact of management practices, and makes this experiment – which followed employees over the experiment and subsequent firm roll-out – particularly informative.

This experiment thus highlights the extensive learning by both the firm and employees around the adoption of a new management practice like working from home. *Ex ante*, both groups were unsure about its impact, and the 9-month experiment and subsequent roll-out process were essential for their ability to evaluate it. These gradual learning effects are one factor behind the slow adoption of modern management practices, and we see the results as being similar to the adoption process for product innovations, like hybrid seed-corn as emphasized in Griliches' (1957) classic article.

This experiment is the first randomized experiment on working from home. As such, it also provides causal evidence to supplement the prior case-study and survey research. It is also unusual in that it involves a randomized controlled experiment within a large firm. Moreover, we were also granted exceptional access not only to data but also to CTrip management's thinking about the experiment and its results. This was because one of the co-authors of this paper, James Liang – the co-founder and Chairman of CTrip during the experiment – was also a doctoral student at Stanford GSB at the time. As a result, this paper benefited from unusually rich insight into the rollout and adoption of a new management practice in a large, multinational firm.

Of course the experiment involved a particular group of employees – those working in call-centers – who tend to be lower paid with a more measurable outcome (i.e. number of calls made) than those in many other jobs. As such, the direct implications for performance are limited to these types of jobs. But as Figure 1 shows, there are still many millions of US employees working from home in lower-paid jobs, many of whom are in roles with measurable outcomes like telesales and IT support. More generally, we also believe that the results on attrition and promotion have broader applicability – many employees do seem to strongly prefer working from home, but fear this reduces their chances of promotion. Our study also highlights the importance of learning and experimentation around working from home – CTrip's management and more than half their employees appear to have changed their views in light of the experiment.

This paper connects to three strands of literature. First, there is a long literature that links the puzzling dispersion of productivity between firms to differences in management practices (see the literature from Walker (1887) to Leibenstein (1966) to Syversson (2011), Gibbons and Henderson (2012) and Bloom *et al.* (2013)). Our paper suggests that uncertainty about the efficacy of new practices may play a role in the slow diffusion of these practices, including those addressing issues of work-life balance such as WFH. These practices have potentially large effects on measured productivity. For example, based on the methodology that is used to measure productivity using Census data (e.g. Foster, Haltiwanger and Krizan, 2000) CTrip would have experienced a measured productivity increase of 30% after introducing working from home, even before accounting for selection effects, because it increased output while cutting capital and labor inputs.

The second strand of literature is on the adoption of workplace flexibility and work-life balance practices. It is based primarily on case-studies and surveys across firms. These tend to show large positive associations of adoption with lower employee turnover and absenteeism and with higher productivity and profitability.[4] However, these studies are hard to evaluate because of the non-randomized nature of the programs. This is true both in terms of the selection of firms into adopting working-from-home programs and also the selection of employees to work at home.

There is also a connection to the urban economics literature. Reducing the frequency of commuting will reduce vehicle miles travelled, lowering emissions, but also reducing population centrality as people move out to the suburbs (Bento et al. 2005).[5] WFH is also part of the wider impact of IT on firm fragmentation arising from the increasing ease of long-distance communicating (e.g. Rossi-Hansberg et al., 2009 and Glaeser, 2013). CTrip in 2012 was setting up regional offices to employ workers in lower-wage, inland Chinese cities using the same working-from-home technology they used in this experiment.

II. MODELLING WORKING FROM HOME

The impact of allowing working from home on firm profits at CTrip was, in the first instance, driven by four effects:

---

[4] See, for example, see the survey in CEA (2010).
[5] For example, jetBlue allowed home-based call center employees to live up to 3 hours drive from the office as they needed to come into the office only one day per month. As a result, many jetBlue employees chose to live in rural areas outside Salt Lake City, the firm's headquarters.

A) *Hours worked*: The number of hours actually worked from the official shift, as opposed to taken on breaks or leaving early

B) *Productivity:* The number of (quality adjusted) calls completed per hour

C) *Attrition*: The impact on quit rates (which drive hiring and training costs)

D) *Capital*: The impact of capital inputs, through office space and equipment requirements.[6]

To highlight these. we consider the firm's profit per worker as a function of the fraction of the work-week spent at home, denoted $\theta \in [0, 1]$. Employees work H hours per week and produce p (quality adjusted) calls per hour. Calls have unit value of r to the firm. The number of hours worked is the employee's choice: $H = 40 - B$, where B is the time spent on breaks, and the scheduled shift length is 40 hours. The location of work may affect the amount of time the employee chooses actually to work, i.e, take on breaks. On the one hand, not having to commute means the employee need not leave early to run errands, and eating at home alone might lead to shorter lunch breaks, since there is no opportunity to socialize. On the other hand, without a supervisor watching them, employees at home might be more easily distracted to other activities. Productivity p also depends on location. A quieter working environment at home might boost productivity, or the lack of supervisory support might lower it. In any case, we treat this relationship of p on θ as exogenous, not a matter of employee choice.

The firm incurs three sets of costs. First, it must pay the employee a fixed base salary, f, and a per call piece rate, w. We take these as given for the analysis, although they obviously might be altered in response to the effects of working from home. Second, if the employee quits, the firm incurs recruitment and training costs of t. We denote the probability of quitting by a, which we take to be a decreasing function of θ: at least for those who want to work at home, being able to do so should make CTrip a more attractive employer. Finally, the firm needs to cover the cost of its real estate and capital equipment, k, at rental rate r. Office space is quite expensive, and WFH four days a week allows an 80% savings on space by "hot-desking." Thus, even though WFH involves having extra computers in the employees' homes, we take k as decreasing in θ.

This yields the call-center profits per employee

$$\Pi = rHp(\theta) - f - wHp(\theta) - ta(\theta) - ck(\theta),$$

where H is determined by the worker's choice and thus may depend on θ.

If the firm considers allowing WFH (setting θ at a positive level), the first order effect on profits is given by

$$(r-w)(H'p + p'H) - ta' - rk'.$$

The last two terms are positive, reflecting the attrition and capital effects. These push for allowing WFH. The term (r-w) is positive, so the focus must be on the terms $H'p$ and $p'H$, which embody the hours worked and productivity effects respectively. If together these are sufficiently negative, as the managers feared, then WFH is not a good idea for the firm.

As noted, we take the dependence of productivity on work location $p'(\theta)$, as exogenous. It might be either positive or negative.

The dependence of employee's effort choice on θ can be modelled simply. Suppose the employee values income, breaks from the mandated work week and leisure time, and that utility may also depend on the location of work, at least in the sense of its altering the relative desirability of the direct arguments: $U = U(Y, B, L; \theta)$.

Recall that income is $f + wpH$, while $H = 40 - B$, where 40 hours is the standard work week. Leisure can be written as $80 - (1 - \theta)c$, where c is the weekly commute time faced by the employee not working at home.

Then utility is
$$U(Y, B, L; \theta) = U(f + wp(\theta)(40 - B), B, 80 - (1 - \theta)B; \theta),$$

where 80 is the number of leisure hours if the employee commutes every day.

The dependence of B on θ is determined by the sign of the cross partial of U with respect to B and θ: if the cross partial is positive, then the employee will take more break time and work less when working from home, and the opposite it the cross partial is negative. If the sign is indeterminate or changes, then no unambiguous prediction can be made.

In this case, the derivative with respect to B is simply $-p(\theta)wU_1 + U_2$, while the cross partial is
$$-wp'U_1 - pw[(40 - B)pw'U_{11} + cU_{13} + U_{14}] + (40 - B)wp'U_{21} + cU_{23} + U_{24}.$$

There are no obvious conditions that will allow signing this unambiguously. If utility is linear in income, however, this simplifies to
$$-wp' + cU_{23} + U_{24}.$$

If, as was found to be the case at CTrip, productivity is higher at home ($p' > 0$), then the first term is negative: Having more leisure directly leads to working more. It seems natural that U_{23} should be negative. Breaks and leisure are probably be substitutes, so more leisure (from not commuting) lowers the marginal utility of taking breaks from work. In this case, more leisure leads to fewer breaks.

The term U_{24} can be interpreted as capturing the effect of location on the utility of breaks. It could go either way, and its relative magnitude is hard to forecast. If the opportunity to socialize makes breaks at work more attractive, or if the need to go outside for a cigarette requires taking a longer break, this would have a negative sign and WFH would lead to fewer breaks and more hours

worked. If reduced monitoring or more distractions at home made breaks at home more attractive, then the sign is positive.

In sum, if productivity is higher at home, leisure and breaks from work are substitutes, and the marginal utility of breaks is lower at when working at home, then the predicted effect of WFH on hours worked is positive. Note that this prediction is not what the management expected. Then given that p' and H' are positive, some measure of WFH is optimal. However, these sign restrictions and the linear utility assumption are not unquestionable, and the actual effects are generally ambiguous.[7] In particular, if working from home makes shirking (from home) more attractive, then H might easily decline under WFH. So what to expect is not at all clear. Not surprisingly, we found see a large number of employees at CTrip changed their minds on working from home over the course of the experiment.

~~The first term, $\partial p/\partial\theta{\times}C$, is the *quality effect*. The sign of this is ambiguous although CTrip's ex ante beliefs were it would be negative – they feared employees would be distracted at home and telephone call quality would drop. This is a common concern in the case-study literature on working from home, although this need not be true if, for example, quieter home conditions improve concentration.~~

~~The second term is, $(p{-}w){\times}\partial C/\partial\theta$, is the *output effect*. The sign of this is again ambiguous[8]. Employees may work longer hours at home if, they take less breaks during work hours as they have more free time outside of work from less commuting. For example, they may go shopping after work when they otherwise would be commuting. Alternatively, this term may be negative if distractions like friends and television increase the marginal value of breaks.~~

~~The third term is $-t{\times}\partial A/\partial\theta$, the *attrition effect*. This was positive for CTrip. Offering employees the option to work from home should (weakly) increase the attractiveness of their job, although it is possible this rises if, for example working from home reduces employees attachment to colleagues.~~

~~Finally, $-r{\times}\partial K/\partial\theta$, is the *capital effect*. This was also positive for CTrip. Having employees working from home for 4 days a week reduced the need for office space by 80% (since employees rotated days in the office and hot-desked). There was a second capital term, which was the increased cost of equipment. The employees' computers and telephones employees they had at home were redundant on the day they came into the office, raising equipment needs by 20%. But for CTrip equipment capital per employee was about \$5k compared to \$24k for office space, so the net effect was positive. However, for firms like manufacturers with large per employee equipment costs this *capital effect* could be strongly negative, making working from home effectively impossible.~~

---

[7] Note that this analysis ignores the possibility for the employee of winning promotion. Promotion was merit-based at CTrip and involved henceforth working in the office. The effects of WFH on the incentives to perform at high levels are ambiguous. Allowing WFH might have discouraged those who really liked working from home from performing at as high as they otherwise would, promotion would involve having to go to the office. At the same time, employees who really wanted promotion and were working from home might work harder to offset the "out of sight, out of mind" effect (although after the experiment ended they could simply return to the office).

[8] Note we assume that p>c since employees are paid a fixed wage in addition to their piece-rate.

~~**Employees**~~

~~While we focus on firms decisions we also consider the factors driving employees to work from home. We assume employees derive utility from income (Y), breaks (B), leisure time (L) and promotions (P). Each one of these will be influenced by working from home.~~

~~Starting with income, this comes from base salary plus performance pay, $Y=f+w\gamma(\theta)(40-B(\theta))$, where $\gamma(\theta)$ is output per hour and $(40-B(\theta))$ is the number of hours worked in terms of the weekly shift less breaks. We have allowed both output and breaks to depend on the location of work, so that the income effect, $\partial Y/\partial\theta$, is ambiguous. Turning to breaks, $\partial B/\partial\theta$ will also be ambiguous, as the impact of working from home will be dependent on the marginal value of breaks. This may be higher at home (e.g. spending time with your wife or children) or at work (e.g. socializing with colleagues). Third, leisure time will be increased by working from home because of the time saved commuting, so $\partial L/\partial\theta>0$. Finally, the impact of promotions is also ambiguous. Negative effects will arise if spending time in the office is important for building human capital (from managers training and feedback) or for office visibility. It could be positive if the impact of working from home on output and work quality is positive, since promotions are functions of employees performance.~~

~~So the net impact of employees utility is ambiguous and complex. Not surprisingly, we found see a large number of employees at CTrip changed their minds on working from home.~~
.

III. THE EXPERIMENT

III.A. The Company
Our experiment took place at CTrip International Corporation, a leading travel agency in China with operations also in Hong Kong and Taiwan. Like other travel agencies internationally, CTrip aggregated information on hotels, flights and tours; made reservations and obtained tickets for clients; and generated revenue through commissions from hotels, airlines and tour operators. Because of lower Internet penetration in China, CTrip did much more of its business on the telephone than leading U.S. agencies like Expedia, Orbitz or Travelocity. CTrip was established in 1999, was quoted on NASDAQ in 2003, and was worth about $5bn at the time of the experiment. It was the largest travel agent in China in terms of hotel nights and airline tickets booked, with over 50% market share in 2010. Exhibit A displays photos of the CTrip Shanghai office, a modern multi-story building that housed the call center in which the experiment took place, as well as several other CTrip divisions and its top management team. The firm also operated a second call center in Nan Tong, a city about 1 hour away from Shanghai, which employed about 5,000 call center staff. Both locations operate in the same fashion, with the same equipment under the same procedures.

Call center representatives were organized into small "teams"[9] of around 10 to 15 people (mean of 11.7 and median of 11), grouped by department and the type of work. There were four jobs in each of the two departments – hotel and airline – involved in the experiment. These were order takers, who answered customer calls, took orders, and entered them into the CTrip information system;

---

[9] The call-center jobs did not involve "teamwork" and there was no group-based pay, but we stick with the term "team" because that was what CTrip called the work groups operating under a common "team" manager.

order placers, who dealt with the airlines and hotels and then notified the clients; order correctors, who resolved problems such as a flight being canceled; plus a night shift that both placed and corrected orders. The members of a given team sat together in one area of the floor, typically occupying an entire aisle. Each call center representative worked in a cubical with equipment including a computer, a telephone and a headset. When team members were ready to start work, they logged on to CTrip's IT system and, in the case of order takers, client calls were automatically routed into their headsets. Order placers and order correctors were also allocated tasks automatically. The allocations between the two Shanghai and Nan Tong call centers were handled centrally, using a standard call queuing system. When employees wanted to take a break, they logged out of the system. The team leaders patrolled the aisles to monitor employees' performance as well as help resolve issues with reservations, provide ongoing training, and give emotional support to employees dealing with difficult clients.

The employees typically worked 5 shifts a week, scheduled by the firm in advance. All members of a team worked on the same schedule, so individuals could not choose their shifts. A team shared the same team leader and the same work schedule, as well as the same working area on the call center floor.

Monthly earnings were composed of a flat wage and a bonus. The flat wage depended on seniority, education, and prior experience, averaging around ¥1300 per month. The bonus portion depended on the individual's monthly performance, and also averaged about ¥1300 per month. The bonus was primarily a linear function of call and order volumes, but with small adjustments for call quality (penalties were applied for call quality scores below certain thresholds) and shift type (night shifts, for example, were paid a higher flat rate). Promotion to team-leader was also based on performance, so both current pay and career concerns provided incentives for employees to perform well.

Since no other Chinese firm had tried allowing home-working among its call center employees, there was no local precedent. In the US, the decision to allow employees in call centers to work from home varied across firms, even those within the same industry, suggesting a lack of any consensus on its impact. Meanwhile, the prior academic literature on call centers also offered limited guidance, being based on case studies of individual, firm-level interventions. Given this uncertainty, management decided to run an experiment.

III.B. The Experimental Design
The experiment took place in the airfare and hotel booking departments in the Shanghai call center. The treatment in our experiment was to work 4 shifts a week at home and to work the 5th shift in the office on a fixed day of the week determined by the firm. Treatment employees still worked on the same schedule as their teammates because they had to work under the supervision of the team leader (who was always office-based), but they operated from home for 4 of their 5 shifts. For example, in a team the treatment employees might work from home from 9am to 5pm on Monday, Tuesday, Wednesday and Friday and in the office from 9am to 5pm on Thursday. The control employees from that team would work in the office from 9am to 5pm on all five days. Hence, the experiment changed only the location of work, not the type or the hours of work. Because all incoming phone calls and work orders were distributed by central servers, the work flow was also identical between office and home locations. Home workers also used the same CTrip-provided computer terminals, communications equipment and software, faced the same pay structure and

undertook the same training as the control group (although for the treatment employees this occurred only the day they were in the office).

Importantly, individual employees were not allowed to work overtime outside their team shift, because doing so would require their team leader to supervise their work. Hence, entire teams could have their hours changed – for example all teams had their shifts increased during the week before Chinese New Year – but individuals were not able to work overtime on their own. In particular, eliminating commuting time, which was 80 minutes a day for the average employee, did not permit the treatment group to work overtime, and so this is not a factor directly driving the results.

Three factors other than location did differ between treatment and control. First, the treatment group's spending less time commuting meant that they would sometimes be able to take care of personal and family responsibilities without taking breaks or leaving early from work. As we will see, this appears to have had a significant effect. Second, the treatment workers did not have as much support from their team leaders, because technological limitations meant that they could not get real-time help from their supervisors while dealing with clients. If anything, this presumably reduced the effectiveness of the treatment workers. Finally, the work environment differed between treatment and control. The former were working alone, typically in what was reported to be a quieter environment. Being alone had some negative effects on willingness to work from home, but the quietness had positive effects on productivity.

In early November 2010, employees in the airfare and hotel booking departments were informed of the working from home program. They all took an extensive survey documenting demographics, working conditions and their willingness to join the program. When asked about their willingness to join the program, employees were not told the criteria that they would have to meet in order to participate in the program. Employees who were both willing and qualified to join the program were recruited for the experiment. Of the 994 employees in the airfare and hotel booking departments, 503 (51%) volunteered for the experiment. They tend to have a longer commute, less tenure in the firm, less education and their own bedroom (see Table 1). Importantly, prior performance (measured simply by the gross-wage, given that almost 50% of earnings was performance-related pay) was not predictive for the take-up of working from home. This helped to assuage one concern of the firm, that worse performing employees would be more tempted to work from home to avoid the direct supervision of their team leaders.

Interestingly, 49% of employees did not opt to work from home, despite the considerable saving in commuting time and cost. The major reason given for this in later interviews was the loneliness of working from home and the lack of opportunities to socialize in the office and after work. Another reason was the possible negative impact for WFH on promotion, which as we discuss in section IV below, appears to have been a legitimate concern.

To qualify to work from home, an employee also needed to have tenure of at least 6 months, have broadband Internet at home to connect to the network, and an independent workspace at home during their shift (such as their own bedroom). Among the volunteers, 249 (50%) of the employees met the eligibility requirements and were recruited into the experiment.

The treatment and control groups were then determined from this group of 249 employees through a public lottery. Employees with an even birth date (a day ending 2, 4, 6, 8, etc.) were selected into the treatment and those with an odd birth date were in the control group. This selection of even birthdates into the treatment group was randomly determined by the Chairman, James Liang, by drawing a ping pong ball from an urn in a public ceremony one week prior to the experiment's start date (see Exhibit B).[10] The treatment group was notified and equipment was installed at each treatment participant's home the following week. Comparison between treatment and control groups (see Appendix A2) shows there was a significant difference between them in 2 of the 18 characteristics: the control group was more likely to be married and have children. This reflects the fact that with 18 different characteristics on average we expect 1 difference to be significant, with 2 differences not unusual. Importantly, when we interact these differences with the impact on WFH we find no impact (Appendix table A5), suggesting this sampling variation is not material for the results.

The experiment commenced on December 6, 2010 and lasted for 9 months, which employees knew in advance. All treatment employees had to remain at home for this period, even if they changed their minds and wanted to return to the office. On August 15, 2011, employees were notified that the experiment had been a success and that CTrip would roll out the experiment to those who were qualified and wanted to work at home in the airfare and hotels departments on September 1st, 2011.

Throughout the experiment, employees were told the experiment would be evaluated to guide future company policies, but they did not learn the actual policy roll-out decision until August 15th. Because of the large scale of the experiment and the lack of dissemination of experimental results beyond the management team, prior to the roll-out decision, employees were uncertain about what that decision would be. Employees in the treatment group who wished to come back to work in the office full-time were only allowed to do so after September 1st, while control group had to stay in the office for the full duration of the experiment.[11] Hence, the treatment and control assignments were fixed for the full 9 months, except for a few cases where equipment broke down or employees changed apartments and lost access to their own room.[12]

Figure 2 shows compliance with the experiment throughout the experimental period, and for one year after the general roll-out. During the experiment, the percentage of treatment group working at home hovered between 80% and 90%. Since compliance was not perfect, our estimators take even birth date status as the treatment status, so we estimate an intention to treat result on the eligible volunteers. The impact on performance if compliance were perfect would be about 10% larger.

After the experiment, we see about 50% of the treatment group immediately decided to return to the office, despite having to incur the financial and time costs of commuting. Strikingly, only about

---

[10] It was important to have this draw in an open ceremony so that managers and employees could not complain of "favoritism" in the randomization process. The choice of odd/even birthdate was made to ensure the randomization was straightforward and transparent.

[11] Employees who would like to work from home in the control group were allowed to work from home after equipment were installed at home starting late October 2011.

[12] In all estimations, we use the even birthdate as the indicator for working-at-home, so these individuals are treated as home workers. In a probit for actually working from home during the experiment, none of the observables are significant, suggesting that returning to the office was effectively random. One reason is that the IT group policed this heavily to prevent employees fabricating stories to enable them to return to the office.

35% of the control employees – who all initially were interested in working from home – actually moved home when they were allowed to do so. The main reasons both groups gave in interviews for changing their minds were concerns over being isolated at home. Finally, we also see that about 10% of the workers who did not initially volunteer changed their minds after the experiment and decided to work from home.

It is worth noting that the firm's management was surprised by two of the findings. First, they were struck by how many employees changed their minds about working from home. More than 50% of the volunteer group and 10% of the non-volunteer group switched preferences after the 9-month experiment. Employees reported that after working from home for a few months they became lonely and wanted to return to the office.

Second, despite the time and financial savings from not having to commute, more than half of the workers eligible to work at home decided to work in the office, suggesting they place a high value on social interactions at work (Hamermesh, 1990). This is particularly striking because, as we note below, we find no negative impact of home working on performance or quality of service.

III.C. Data Collection

CTrip had an extremely comprehensive central data collection system, in large part because its founding team came from Oracle and had extensive database software experience. The bulk of the data we used in our paper were directly extracted from the firm's central database, providing extremely high data accuracy. The data we collected can be categorized in 7 fields: performance, labor supply, attrition, promotions, reported employee work satisfaction, detailed demographic information, and survey information on attitudes towards the program.

Performance measures varied by the job: the 134 order takers versus the 11 order placers, order correctors, and night shift workers (details in Appendix 1). Order takers' key performance measures were the number of phone calls answered and number of orders taken. The key measures for the other three groups were the numbers of different types of calls made. For order takers, we could also accurately measure time spent working (in terms of minutes on the phone) because phone calls and call lengths were logged in the central CTrip database. The firm used these measures to monitor the work of its employees. We also calculated phone calls answered per minute as a measure of labor productivity for these workers.

We have daily performance measures of all employees in the airfare and hotel booking departments from January 1st, 2010 onwards as well as daily minutes on the phone for order takers. We also collected data on different types of promotion by September 2012, almost 2 years after the experiment commenced. The firm ran internal surveys of the employees during the experiment on work exhaustion and on positive and negative attitudes (see details in Appendix A3). We conducted two rounds of surveys, in November 2010 and August 2011, to collect detailed information on all the employees in the two departments including basic demographics, income, and attitudes toward the program, and ran extensive interviews and focus groups with the employees. Finally, in May 2012 we conducted interviews with managers and with employees who had chosen not to work from home to learn about the reasons for this choice.

IV. IMPACT ON THE FIRM

We analyzed the effect of home-working both in terms of its impact on the firm, which we cover in this section, and the impact on the employees, which we cover in the next section. The data and Stata do-files to replicate all results are available on www.stanford.edu/~nbloom/WFH.zip.

IV.A. Individual Employee Performance

We first estimated the intention to treat effect on weekly employee performance for the eligible volunteers prior to and during experimental period data via equation (1)

$$Employee\ Performance_{i,t} = \alpha Treat_i \times Experiment_t + \beta_t + \gamma_i + \epsilon_{i,t} \quad (1)$$

where $Treat_i$ is a dummy variable that equals 1 if an individual belongs to the treatment group defined by having an even-numbered birthday; $Experiment_t$ is a dummy variable that equals 1 for the experimental period December 6th to August 14th; and $Employee\ Performance_{i,t}$ is one of the key measures of work performance. This includes the log of weekly phone calls answered, log of phone calls answered per minute on the phone, log of weekly sum of minutes on the phone, and an overall performance z-score measure (performance score normalized to mean=0 and standard-deviation=1 based on pre-experiment performance for each task). Finally, β_t reflects a full set of weekly time dummies to account for seasonal variation in travel demand, such as the World Expo in 2010 and the Chinese New Year, and γ_i reflect a full set of individual fixed effects.

To make performance of different types of workers comparable, we use performance z-scores. For each individual we subtract the pre-experiment mean for the set of individuals holding the same job ("worker type") and divide by the pre-experiment standard deviation for the worker type. Hence, this normalized z-score measure has a mean 0 and standard deviation 1 across all employees within each type of worker during the pre-experiment period. We do this both for individual performance measures and for the pair of measures for the order takers, in which case the z-scores on the individual measures are used to create an overall z-score.

In column (1) of Table 2, overall performance of the treatment group is found to be 0.232 standard deviations higher than the control group after the experiment started, significant at the 1% level. Column (2) limits the sample to only performance data during the experiment and compares the treatment group to the control group without controlling for individual fixed effects. The coefficient is slightly smaller but similar.[13] If we limit the sample to the 134 order-takers, we can use phone calls answered as the key performance measure. The z-scores of phone calls account for different volume and average length of phone calls in two departments. In column (3), we look only at the phone calls performance measure and find it is 0.248 standard deviations higher in the treatment group. In column (4), we look at the log of phone calls and find a coefficient of 0.120, meaning that treatment employees were making 13% (noting that 13%=exp(0.120)) more phone calls per week.

---

[13] Because we have a randomized intervention we can examine either the *difference* between treatment and control (evaluated over the experimental period), or the *difference of differences* (evaluated as the change in performance between treatment and control over the experimental period versus the pre-experimental period). Since employees have large pre-existing cross-sectional variations in performance we obtain more accurate estimations from using the difference in difference specification, estimated using the panel with employee fixed effects. However, comparing columns (1) and (2) we see the estimators are quantitatively similar and within one-standard deviation of each other.

We can also see these results in Figure 3 where we plot the raw number of phone calls per week for the treatment and control groups from Jan 1st 2010 until the end of the experiment in August 14th 2011. Before the experiment started, the treatment group trended closely together with the control group, both of which bounced around due to seasonal fluctuations in demand. But once the experiment began, the treatment group started to outperform the control group, answering about 40 more phone calls per person per week.

Interestingly, the difference in performance was greatest during the middle of the experiment, from about 2 to 6 months. The smaller rise in improvement during the first 2 months was due to installation and learning effects. It took several weeks for all the IT and logistical bugs to be addressed. The gradual decline in the performance gap from 6 months onwards reflects two trends. First, poorly performing employees in the control group were more likely to quit than those in the treatment group (see section IV.B and Table 8), boosting the control group's performance. Second, from focus groups and interviews we learned that some employees in the treatment group felt lonely working at home after a few months and wanted to return to the office but could not because of the experiment's design. This potentially affected their motivation.

Figure 4 plots the cross-sectional distribution of performance for treatment and control groups at 3 months into the experiment, displaying a broad distributional improvement from working-from-home (rather than the results being driven by a few outliers).

We further decomposed the difference in performance observed in column (4) into phone calls answered per minute on the phone (a measure of productivity), and minutes on the phone (a measure of high-frequency labor supply). In column (5), we found treatment employees were making 3.3% (note that 3.3%=exp(0.032)) more phone calls per minute, which the employees attributed to home being quieter than the office. They told us this meant it was easier to hear the customers, so they did not have to ask them to repeat themselves as often and could process the orders more quickly. This suggestion matches the psychology literature, which has shown that background office noise can reduce cognitive performance (see, for example, Banbury and Berry, 1998).

The largest factor increasing the home-workers performance is that, as shown in column (6), they worked 9.2% (9.2%=exp(0.088)) more minutes per day. This was despite the fact that home and office workers both worked the same nominal shift. The reason home-workers could increase minutes on the phone was that, within their shifts, they were available to take calls for more time, meaning they were taking less time off during their shifts.

IV.B. Individual Employee Labor Supply
In Table 3, we investigate the factors driving this increase in minutes worked within each shift. Column (1) repeats the results from the final column of Table 2. Columns (2) and (3) break this difference in minutes on the phone down into two pieces. In column (2), we look at whether treatment workers spent more minutes on the phone per day worked, and column (3) looks at whether they worked for more days.

Column (2) shows that about three quarters of the difference in the time on the phone was accounted for by the treatment group's spending more time on the phone per day worked. This is because: (a) they started work more punctually, a phenomenon which they attributed to avoiding the effects of events that disrupted commuting like bad traffic or the heavy snow in Shanghai in February 2011;[14] (b) they could schedule personal matters, like doctor's appointments, in the time they saved by not commuting (rather than having to leave early); and (c) they took shorter breaks during the day. In column (3) we see that the other one-quarter of the difference in time worked between treatment and control was explained by treatment employees' working more days because they also took fewer sick-days. Employees explained this was because they continued to work at home when they felt somewhat ill but would not have felt up to commuting into work.

To investigate this further we interacted the WFH treatment dummy with a long commute indicator. Individuals with long commutes should be more likely to increase their hours when WFH since they are more likely to suffer commuting delays. Indeed looking at columns (4) to (6) we see this: the differences in the weekly time on the phone as well as time on the phone per day worked are larger for employees with a commute times of above 120 minutes per day.[15] We also tried other interactions with marital status, children, education and tenure (the variables that predicted WFH in Table 1) and found no significant interactions.

IV.C. Quality, Spillovers

One question is whether quality of the service was compromised for the increase in output in the treatment group. We constructed two quality measures: conversion rates and weekly recording scores. Conversion rates were calculated as the percentage of phone calls answered that resulted in orders, while the weekly recording scores came from the 1% of phone-calls that were randomly evaluated by an external monitoring team. In summary (with the full details in table A4 in the appendix), we find no impact of working from home on quality using either measure.

Another related question is whether the improvement associated with working from home came from an improvement in the treatment group or from a deterioration in the control group. Perhaps the gap between treatment and control was caused not by the treatment group performing better but by the control group performing worse after they "lost" the randomization lottery. The group winning the treatment lottery saved themselves 9 months of commuting time and costs, a substantial gain worth about 17% of their salary, evaluated at their CTrip wage rate.[16]

We collected data on two other "quasi" control groups to answer this question. The first group was the employees in the Nan Tong call center who would have be eligible under the selection rules for the experiment. This call center also had airfare and hotel departments, and calls were allocated across the Shanghai and Nan Tong call centers randomly from the same central server. The second group was the 190 employees in the Shanghai call center who did not volunteer to participate in the WFH experiment but met the eligibility requirements to work from home. We think these two

---

[14] CTrip is strict with punctuality of its workers. If a worker comes to work late, it is recorded as taking time off.

[15] 21.3% of employees commute for more than 120 minutes per day (the includes the commute to and from work).

[16] The average employee made about $100 per week for a 40 hour week. The average commuting time was 40 minutes each way and the corresponding out-of-pocket commuting cost averaged $0.50. Hence, the saving in time is about $13 a week in time costs and about $4 per week in out–of-pocket costs.

groups were comparable to the treatment and control groups for two reasons. First, all four groups faced the same demand for their services. Second, they all met the requirements for eligibility to participate in the experiment.

Comparisons of these alternative control groups are also reported in Table 4. Results in the top panel of Table 4 compare the Shanghai treatment and control groups to Nan Tong, revealing no significant differences in overall performance between Nan Tong and the control group. The bottom panel compares treatment and control group to the eligible non-experimental group in Shanghai. Again, we find no significant difference between the control group and the eligible non-experimental group. These results suggest that the gap between the treatment and control group reflects an improvement in the performance of the treatment group rather than any deterioration of the control group. That is, although the control group and the treatment group work in the same team, we find – perhaps surprisingly – no evidence of the control group's having been discouraged by losing the working-from-home lottery. [17]

IV.D. Potential Hawthorne Effects
Another explanation for the superior performance of the treatment group are Hawthorne effects, which suggest that the employees were motivated by the experiment, possibly deliberately so that the firm would roll out WFH permanently. We should note three things, however, that make this appear unlikely. First, there were 131 employees working from home, so each individual employee had little impact on the overall evaluation of the experiment, and thus little incentive to try to manipulate it. Second, the gap between home-based employees and office-based employees widened after the experiment ended, as we shall see in the next section. Finally, the firm was itself so convinced that the success of the experiment was not due to Hawthorne type effects that it rolled out WFH to the entire airfare and hotel divisions.

IV.E. Post-Experiment Selection
In August 2011, the management estimated that each working from home employee saved CTrip about \$2,000,[18] so they decided to extend the option to work from home to the entire hotel and airfare departments. Employees in these departments were notified that the experiment had ended and they were entitled to choose their location of work (conditional on being eligible), so control employees who still wished to could move home, and treatment employees who wanted to return to the office could do so.[19]

As shown in Figure 5 – which plots the difference in normalized phone-calls between home and office workers – post-experiment selection substantially increased the performance impact from working from home. The reason is that workers that had performed relatively badly at home returned to the office. This increased the increase in performance z-scores from home-working from about 50 calls per week during the experiment to more than 100 calls per week 9 months after the

---

[17] Alternatively, we could compare the performance of treatment and control group during the days treatment group work at home with the days they work in the office. Unfortunately, they were not comparable because team leader often scheduled weekly team meetings during the day treatment group worked in the office.

[18] See Section V.A for the derivation of this figure.

[19] Treatment group employees who were working at home and wanted to come back to work in the office full-time were allowed to come back on September 1st, 2011, 2 weeks after the announcement of the roll-out. Control group employees who wanted to work at home started to do so gradually from the beginning of November 2011.

end of the experiment. This is also evaluated in Table 5, which estimates the performance impact of working-from-home during and after the experiment.

In Table 5 column (1) repeats our baseline results for z-scores. In column (2) we see that the average z-score rose by 28.5% after the experiment. Once we control for quits by using a balanced panel we find even larger increases in performance, of 34% for z-scores in column (3). In column 4 we examine instead our direct performance measure, which is the number of phone calls again repeating our baseline specification. Column (5) is the key result – the post experiment effect of WFH rises to 24.6%[20], almost double the 13% increase in the baseline. Finally, in column we again control for quits with the balance panel and the increase in performance after the end of the experiment is 105% of the baseline.[21] The fact that the balanced panel results show an even larger increase in performance indicates that it is sorting of employees between home and the office (rather than differential attrition) that drove the further improvement from WFH during the post-experiment, roll-out period.

This sorting effect was driven by treatment workers who had performed relatively badly at home returning to the office. This is shown in Table 6, top panel, columns (1) to (4), which runs probits on whether a treatment worker returned to the office. The results show that treatment workers who performed relatively worse at home versus the office returned to the office. This was despite the fact that all treatment workers had initially volunteered to work from home, suggesting that many of them subsequently discovered home working was not as attractive as they initially believed it would be. In the bottom panel we find no significant evidence of differential sorting in the office-based control group, suggesting less learning occurred in this group who did not get to experiment with home-based working.

These results are similar to those of Lazear (2000) and Bandiera, Barankay and Rasul (2007) who find the introduction of performance related pay improves performance both by motivating individual employees to work harder, but also sorting higher ability employees into work. In our case the selection effect is over the location of work, suggesting the impact of many management practice changes may have these additional selection effects if they are introduced as voluntary.

V. IMPACT ON THE EMPLOYEES

V.A. Employees' self-reported outcomes

CTrip management was also interested in how employee self-reported wellbeing was affected by the program. They thus ran two sets of surveys: a satisfaction survey and a work attitude survey. Details of survey questions and methodology are listed in Appendix A3. In summary, these were standard employee satisfaction tests developed by Berkeley psychologists Christina Maslach and Susan Jackson in the 1970s. The satisfaction survey was conducted five times throughout the experimental period: once in early November, before the randomization took place, and four times after the experiment had started. The first three columns of Table 7 show three different satisfaction

---

[20] 24.6%= exp(0.220)-1
[21] 105%=[(exp(0.203)-1)/(exp(0.104)-1)]-1

measures. The treatment group reported no difference in satisfaction levels from the control group at the first survey, but the treatment group reported statistically significantly higher satisfaction once the experiment began.

The work attitude survey was conducted every week. The first weekly survey was conducted in late November 2010, before the experiment began but after the randomization had occurred. Interestingly, the treatment group already reported higher positive attitude (significant at the 10% level), less negative attitude and less exhaustion from work. This group had yet to move home, so this difference was presumably due to the control group's learning they lost the working from home lottery while the treatment group learned they had won. This highlights the importance of comparing our treatment group with other control groups like Nan Tong and the eligible non-experiment group. After starting the experiment, the gap between the treatment and control group rose further, so that the treatment group reported statistically significantly higher positive attitude and less work exhaustion.

V.B. Attrition
One of the key initial reasons CTrip was interested in running the experiment was to lower their rate of employee attrition. The turnover rate among CTrip call center representatives had historically hovered around 50% per year, which was typical of the call center industry in China.[22] Training a new worker cost about six weeks salary, and there were also costs of identifying, screening and hiring new employees. Figure 6 plots the cumulative attrition rate of treatment and control group separately over the experimental period. Shortly after the commencement of the experiment, cumulative attrition rates diverged between the two groups and the difference was statistically significant. By the end of the experiment, the total attrition rate in the treatment group (17%) was less than half of that in the control group (35%) during the 9 months period.[23]

We further tested whether selective attrition existed by running probit regressions in Table 8. The dependent variable is whether an employee quit the job during the experimental period between December 6, 2010 and August 14, 2011. Column (1) confirms the finding in Figure 6 that treatment employees' rate of attrition was about half that of the control group. In column (2), we tested whether employees with worse performance were more likely to leave the firm from the treatment group compared to the control group. To measure performance we used the average individual weekly performance z-scores during the experimental period (from December 6th 2010 until August 14th 2011) or, for employees who quit, their performance until their last full week. We found that low performers were significantly more likely to quit, particularly in the control group.

In columns (3) and (4), we estimated the impact of experimental period performance on quitting in the treatment and control groups separately. We found a large and strongly significant impact in the control group and smaller, but still weakly significant impact in the treatment group. Interviewing the employees, we heard that control group employees who underperformed tended to quit for other similar call center or office jobs, which were easy to find. Treatment employees, however, were

---

[22] 2010 Report on Chinese Call Center Operation and Management. Note that CTrip could in principle fire employees, but this was rare and no employees in these two divisions were fired over this period (as far as we are aware).
[23] These figures are dependent on the market circumstance: the fact that no other call centers offered home-working in Shanghai was likely to CTrip's practice particularly effective in reducing attrition.

much less likely to quit because no other comparable home-working jobs existed, substantially reducing selection from the treatment group.

This differential attrition, of course, also raises the question of whether our estimated impact of WFH is biased. Specifically, the concern is that the estimated impact is mainly driven by differential attrition. We note that this is unlikely, because in Table 7 we see that employees with worse performance in the control group were more likely to quit. This suggests that the observed control group performance level was larger than it would have been, hence generating a smaller performance gap between treatment and control than if no attrition had occurred. In other words, our estimated treatment effect is actually biased downwards.[24]

To address this issue more formally, we used the Lee (2008) bounds estimator. This provides upper and lower bounds on the effect of differential selection on performance across groups, assuming that attrition is monotonically driven by the performance variable. This allows us to generate two bounds – the upper bound that assumes that the extra attrition in the control group is based on a negative correlation with performance (as we saw in Table 7), while the lower bound assumes a positive correlation of attrition with performance (the reverse of what we see in Table 7, but included for completeness). These bounds are shown in Online Appendix Figure A1. We see that the upper bound – which is the most plausible given that is assumes low performers are more likely to quit as found in Table 7 - lies above the actual treatment-control estimated impact. This suggests that the actual treatment effect is if anything larger than we estimated, because the attrition of the worst performers from the control group biases our results down. Hence, even our large estimated positive impact of WFH is likely to be underestimated by attrition bias.

V.C. Promotions and career concerns
One possible negative effect from working at home is that long-run career opportunities could be damaged by less on-the-job training from team leaders and less "face-time" in the office. To investigate this, we collected promotion data on the 249-employee experimental sample. We defined promotion as either being promoted to team leader or being moved to a more advanced function (i.e., switch to the quality-control team), which often occurred shortly before promotion to team leader. Both would require a formal evaluation by team leaders and departmental managers. During the period from the start of the experiment in December 2010 until the end of September 2012, a total of 17 employees from the treatment group received promotions or more advanced job functions and 23 from the control group.

In Table 9 we show the results from running probit regressions on performance before the experiment, performance during the experiment, as well as some demographic controls. We see in column (1) that working from home has no overall impact on promotion. However, we see in column (2) that once we controlled for their performance during the experimental period WFH had a negative and weakly significant impact on promotion. It appears that WFH improved performance – increasing the probability of promotion – but this was offset by a roughly equal negative effect from being home based. In column (3) we add demographics and the story is very similar. Finally, in column (4) we interact WFH with the performance measure and find a negative coefficient in addition to a negative WFH levels effect.

---

[24] Of course, to the extent that lower-performing employees quit, the firm is less concerned with their loss, although the firm still would have rather kept them because of the substantial costs of recruitment and training.

One story that is consistent with this is that home-based employees are "out of sight, out of mind." As a result supervisors did not notice their performance as much and were less likely to promote them. We heard anecdotal evidence for this from employees and managers during focus groups and interviews, and it was one factor that led some employees to return to the office to avoid what they perceived as a WFH promotion 'discrimination' penalty. Another possibility is that WFH employees lack opportunities to develop the interpersonal skills to succeed in managerial jobs and therefore are less likely to be promoted. Either way, it is clear that WFH comes with a promotion penalty.

VI. PROFIT, PRODUCTIVITY AND FIRM LEARNING

One of the most interesting aspects of the experiment was the learning process for both the firm and the individual employees on the costs and benefits of working from home. Both groups were initially unsure about these, because no other Chinese call center had ever offered this option before. However, we were able to monitor both management's and employees' learning over the course of the experiment because of our extensive access to the CTrip's management team and frequent employee surveys and interviews. Before discussing this we first present the estimated impacts on firm profits and productivity from allowing employees to work from home.

VI.A. Profit and productivity impact

The firm saw working from home as a way to save on office costs, but was worried that employees would shirk at home or that call quality would decline due to multi-tasking on other activities which are prohibited in the office, like playing computer games or watching TV. While managers had previously been allowed to work from home on an ad hoc basis, no non-managerial level employees had been allowed to home work. The research literature provided very little guidance on what might happen, with what little evidence there was suggesting that typically routine jobs were, if anything, less effectively carried at home.[25]

Running the experiment revealed, however, that working from home actually generated an improvement in employee performance, worth about $375 per employee per year (evaluated at the 13% performance improvement from the Table 3). In addition, the firm estimated office cost savings of about $1250 per employee and reduced turnover savings of about $400 per employee per year. Hence, given the annual saving of about $2000 per employee, the firm rolled the program out in August 2011, accompanied by an aggressive poster campaign to persuade employees to take up the home working option.

---

[25] For example, Dutcher (2012) ran lab experiments on routine and non-routine tasks with and without remote monitoring, and found the more routine ones were negatively affected by mimicking a home-based environment. He conjectured that the lack of peer and manager effects, which that have been shown to be important in low-level tasks in field environments by Falk and Ichino (2006), Bandiera, Barankay and Rasul (2005) and Mas and Moretti (2009), could explain this.

A related question is what was the impact on total factor productivity (TFP)? We estimate TFP would increase by about 30% from moving every employee home, using the methodology adopted on US Census data in, for example, papers like Foster, Haltiwanger and Syverson (2008) and Syverson (2011).

This 30% rise in measured TFP comes from three sources. First, output (as measured by the number phone calls) increased 13% from working from home. While 9.5% of this increase comes from employees' working more hours, this increase in attendance would not be identified in US Census survey data, since the Census collects information on shift-hours (i.e. 40 hours per week), not actual hours worked. Second, the reduction in attrition from 50% to 25% would reduce steady-state labor hours lost to training by 3%, since new employees need 6 weeks of training.[26] Finally, the capital per employee is comprised of about $5k of desktop IT equipment, $10k of central IT equipment (servers and the network) and $24k of office space (total office capital divided by the total number of employees). Moving employees home for 4 days a week reduces the office space required by 80%,[27] although it increases the desktop IT requirement by 20% (equipment lies idle at home for 1 day a week). On aggregate this reduces capital by 48%. Assuming a production function coefficient of 1/3 on capital and 2/3 on labor, this yields an estimated TFP increase of 30%. Given that the cross-sectional standard-deviation of TFP reported in Foster, Haltiwanger and Syversson (2008) is 26%, this highlights how differences in the adoption of WFH across firms could potentially account for quite large differences in measured productivity.

VI.B. Firm learning

The firm learned four important results from running the formal experiment versus the non-randomized pilot that they had initially been considering. First, they learned that WFH improves performance. Without running a formal experiment, their view was that they could have interpreted the drop in treatment performance shown in Figure 3 as a negative treatment effect. The period of the experiment (December 2010 to August 2011) coincided with a business slow-down for CTrip due to a combination of the (predicted) end of Shanghai Expo 2010 and an (unpredicted) increase in competition from other travel agencies. As a result, the *difference* in performance for the treatment group relative to their pre-experiment baseline was negative, and is only positive when evaluated as a *difference of differences* against the control group. This highlights the importance of having a well matched (ideally randomized) control group to strip out these kinds of seasonal and demand effects.

Second, *ex ante* there was very little discussion of selection effects on employee performance, but by running the experiment and then rolling it out it became clear that allowing employee choice generated a far greater effect than requiring working from home. The impact of working from home was positive, on average, but appears to have a large variance, so that employee choice led to a much higher effect, as shown in Figure 5.

---

[26] Training costs 6 weeks of employee time, which given a 50% rate of employee turnover, has to be amortized over 2 years, meaning in steady state about 6% of employee output is lost from training. Hence, reducing attrition by 50% reduces training time by 3%. This ignores the labor savings from a smaller central recruitment team, as it only focuses on the lost output from training.

[27] Note that since employees hot-desk (share desks), and because teams come in on different days of the week, the firm can reduce the numbers of desks in line with the reduction in the number of days spent working in the office.

Third, having the large sample of treatment and control employees allowed the firm to evaluate the impact on different types of employees. Somewhat surprisingly, they found a very homogeneous impact across all types of employees. For example, in Figure 7, we plot the impact on the top half of the treatment versus control distribution and the bottom half of the treatment vs control distribution. To calculate this, both groups were split in half by the pre-experiment median performance and then compared. What we see is a similar improvement in performance for both groups. CTrip's *ex ante* expectation was that the bottom half of employees were the less motivated ones, and they would perform far worse at home. Table A5 shows a similar result, that the impact of WFH was homogeneous across a range of other characteristics, including gender, commute time, age, prior experience and living arrangements. In other words, in the current context observed characteristics are not good predictors of who will benefit more from work from home. These results have led CTrip to offer WFH to all employee groups rather than any selected sub-samples (such as high-performers), which they initially intended to target.

Finally, management were surprised by the dramatic drop in attrition that highlighted how many of their employees valued working-from-home. They anticipated a reduction, but nothing like the 50% drop they observed.

VI.C. Employees' learning
One direct measure of the extent of employee learning is the number of employees who changed their minds about working from home. Figure 2 shows that after the experiment about 50% of the initial treatment and control volunteers changed their minds and decided to work in the office after the end of the experiment, while 10% of the initial non-volunteer group opted to work from home.

We also designed a survey to inquire into employees' evolving views toward the Program from across all 994 airfare and hotel department employees. We administered the same survey with the help of the CTrip management in November 2010 and August 2011. Employees were asked specifically whether they were interested in participating in the Work-at-Home Program if they were eligible. They could choose from three answers: "yes", "no" or "undecided". We find of the 568 employees that took part in both surveys, that only 303 (53%) maintained their views, while the remaining 47% changed their minds. Of those, 24% went from "yes" or "undecided" to "no", while 12% went from "no" or "undecided" to "yes", with the remainder switching from "yes" or "no" into "undecided".

In follow-up interviews, most of the interviewed employees who had decided they no longer wanted to work from home cited social reasons. Another group who had thought working from home would be attractive found that it was troublesome for the people with whom they lived (often parents), especially if they were called to work outside normal business hours. In reverse, a number of employees switched to working from home because they saw the success of their peers who worked from home..

VI.D. Why did the firm not introduce working from home before?
Finally, one question that arises is why CTrip (or any other similar firm) did not introduce working from home earlier, given it was highly profitable? From extensive discussion with the senior management team there seemed to be two reasons.

First, there was the classic free-rider problem that arises with all forms of process innovation where the absence of intellectual property makes it hard to prevent imitation. CTrip believed that the private benefits of home-based working would be short-lived (if it was successful), as rivals would copy the scheme and use it to drive down commission margins in the travel agent market, while the costs of experimentation would be borne entirely by CTrip. Hence, they viewed themselves as paying the full cost of experimentation but only capturing part of the benefits because of imitation based on knowledge spillovers.

Second, within CTrip the senior management had incentives that provided limited upside and extensive downside from the experimental outcomes. Senior managers were primarily motivated by career concerns, with limited bonus or equity compensation. As a result, their incentives to experiment were muted – they gained little from a successful experiment, but risked major career damage if the experiment was unsuccessful. James Liang, the Chairman and co-founder, had more balanced incentives to promote the experiment since he owned extensive CTrip equity and had no firm-level career concerns. He played a major role in persuading the CTrip executives to run the experiment.

Both factors – the threat of imitation and risk-aversion from the career concerns of senior managers – are likely to represent forces deterring process innovations in most large firms. This suggests these could lead to under-experimentation in managerial and operational practices.

VII. CONCLUSIONS

The frequency of working from home has been rising rapidly in the U.S. and Europe, but there is uncertainty and skepticism over the effectiveness of this, highlighted by phrases like "shirking from home". We report the results of the first randomized experiment on working from home, run in a 16,000 employee, NASDAQ-listed Chinese firm, CTrip. Employees who volunteered to work from home were randomized by even/odd birth-date into a treatment group who worked from home four days a week for nine months and a control group who were in the office all five days of the work week. We found a highly significant 13% increase in performance from home-working, of which 9% was from working more minutes of their shift period (fewer breaks and sick days) and 4% from higher performance per minute. We found no negative spillovers onto workers who stayed in the office. Home workers also reported substantially higher work satisfaction and psychological attitude scores, and their job attrition rates fell by over 50%. Further, when the experiment ended and workers were allowed to choose whether to work at home or in the office, selection effects almost doubled the gains in performance.

This experiment highlights how complex the process of learning about new management practices is. For CTrip, having no precedent in terms of similar Chinese firms that had adopted working from home for their employees led them to run this extensive field experiment. Given their success, other firms are now likely to copy this practice, generating the type of gradual adoption of a new management practices that Griliches (1957) highlighted. More generally, given the large impact of this practice on employee performance – a $2000 per employee reduction in costs and a 30% increase in TFP – this also provides a management-practice based explanation for heterogeneous firm performance.

While our results suggest a promising future for working from home, we should note that several distinctive factors at CTrip contributed to the success of the experiment and the implementation of the practice. First, the job of a call center employee is particularly suitable for telecommuting. It requires neither teamwork nor in-person face time. Quantity and quality of performance can be easily quantified and evaluated. The link between effort and performance is direct. These conditions apply to a range of service jobs such as sales, IT support, and secretarial assistance, but they are far from universal. Second, the firm can closely monitor the performance and labor supply of the employees thanks to its extensive centralized database. Team leaders managers could generate a report from the database of the performance of the team members daily and easily detect problems in individual employees' performance. Third, the extent of WFH was limited, so that it did not require a significant reorganization at the workplace. Because the participation rate ranged between 15% and 25%, team leaders continued to supervise their teams with a mix of home and office workers without any major reshuffling of team membership.

While these features arguably favored successful implementation of working from home at CTrip, we believe the practice of working from home is worth further exploration. After all, much of the research for this paper and its writing were done by the authors working from home.

V. BIBLIOGRAPY

Bailey, Diane and Nancy Kurland. 2002. A review of the telework research: findings, new directions and lessons for the study of modern work. *Journal of Organizational Behavior* 23:283 – 400.

Banbury, Simon and Berry, Dianne. (1998). "Disruption of office related tasks by speech and office noise". *British Journal of Psychology*, 89, 494-517.

Bandiera, Oriana, Barankay, Iwan and Imran Rasul (2007), "Incentives for Managers and Inequality Among Workers: Evidence from a Firm Level Experiment", *Quarterly Journal of Economics*, 122(2), 729-773.

Bandiera, Oriana, Barankay, Iwan and Imran Rasul (2005), "Social preferences and the response to incentives: evidence from personnel data", *Quarterly Journal of Economics*, 120(2), 915-962.

Bento, Antonio, Cropper, Maureen, Ahmed Mobarak and K. Vinha (2005). "The Impact of Urban Spatial Structure on Travel Demand in the United States" *Review of Economics and Statistics*, Vol. 87, No. 3pp

Bloom, Nick, Benn Eifert, Aprajit Mahajan, David McKenzie and John Roberts, (2013), "Does Management Matter: Evidence from India,"*Quarterly Journal of Economics,* 1-51.

Bloom, Nick and Van Reenen, John, (2011), "Human resources and management practices", Handbook of Labor Economics, Volume 4, edited by Orley Ashenfelter and David Card.

Bloom, Nick, Tobias Kretschmer and John Van Reenen, (2009), "Work-life Balance, Management Practices and Productivity', in *International Differences in the Business Practice and Productivity of Firms*, Richard Freeman and Kathryn Shaw (eds.). Chicago: University of Chicago Press.

Clark, L and Tellegen, A (1988), "Development and validation of brief measures of positive and negative affect: the PANAS scales.", Journal of Personal Social Psychology. 1988 Jun;54(6):1063-70.

Council of Economic Advisors (2010), "Work-life balance and the economics of workplace flexibility", http://www.whitehouse.gov/files/documents/100331-cea-economics-workplace-flexibility.pdf

Dutcher, Glenn (2012), "The effects of telecommuting on productivity: An experimental examination. The role of dull and creative tasks", *Journal of Economic Behavior and Organization*, April 2012.

EIROnline (European Industrial Relations Observatory on-line) (2010), "Telework in the European Union," http://www.eurofound.europa.eu/eiro/studies/tn0910050s/tn0910050s.htm

Foster, L., Haltiwanger, J. and Krizan, C.J. (2000), "Aggregate Productivity Growth: Lessons from Microeconomic Evidence", New Developments in Productivity Analysis, NBER, University of Chicago Press.

Foster, Lucia, John Haltiwanger and Chad Syversson (2008), "Reallocation, Firm Turnover, and Efficiency: Selection on Productivity or Profitability?" American Economic Review, 98(1), 394-425

Gibbons, Robert, and Rebecca Henderson (2012), "What Do Managers Do? Persistent Performance Differences in Seemingly Similar Enterprises,' in Robert Gibbons and John Roberts, eds. *The Handbook of Organizational Economics*, Princeton, NJ: Princeton University Press.

Glaeser, Edward (2013), "Employees belong in Cities, not just in cubicles", Bloomberg news http://www.bloomberg.com/news/2013-03-06/employees-belong-in-cities-not-just-in-cubicles.html

Griliches, Zvi (1957), "Hybrid Corn: An Exploration in the Economics of Technological Change", Econometrica, volume 25 (4), pp. 501-522.

Hamermesh, Daniel (1990), "Shirking or Productive Schmoozing: Wages and the Allocation of Time at Work," *Industrial and Labor Relations Review*, January

Falk, Armin and Andrea Ichino. (2006). "Clean Evidence on Peer Effects". *Journal of Labor Economics* 24(1), pp:39 – 57

Lazear, Edward, 2000. 'Performance Pay and Productivity', *American Economic Review* 90(5), 1346-1361.

Lee, David (2008), "Training, wages and sample selection: estimating sharp bounds on treatment effects". *Review of Economic* Studies, 76(3), pp. 1071-1102.

Leibenstein, H. (1966) "Allocative Efficiency vs." X-efficiency," *American Economic Review,* 56(3): 392-415

Mas, Alexandre and Enrico Moretti (2009), "Peers at Work", *American Economic Review* 99 no. 1:112 - 145

Maslach, C., & Jackson, S.E. (1981). *Maslach Burnout Inventory*. Research edition. Palo Alto, CA: Consulting Psychologist Press.

Mateyka, Petr J., Melanie Rapino and Liana Christin Landivar. (2012) "Home-Based Workers in the United States: 2010," United States Census Bureau, Current Population Reports.

Oettinger, Gerald (2011), "The Incidence and Wage Consequences of Home-Based Work in the United States, 1980-2000", *Journal of Human Resources,* 46(2): 237-260

Rossi Hansberg, Esteban, Sarte, Pierre-Daniel and Owens, Raymond (2009), "Firm fragmentation and urban patterns", *International Economic Review*, 50(1): 143-186.

Syversson, Chad (2011), "What Determines Productivity at the Micro Level?", *Journal of Economic Literature*, 49(2): 326-365

DATA APPENDIX

The data and Stata do-files to replicate all the results in the paper are available on www.stanford.edu/~nbloom/WFH.zip.

Appendix A1: Table for different types of workers and their key performance measures

Types of Workers	Department	Key Performance Measures	Number of Workers
Order Takers	Airfare	Phone Calls Answered	88
	Hotel	Orders Taken	46
Order Placers	Airfare	Notifications Sent	43
	Hotel	Reservation Phone Calls Made	25
Order Correctors	Hotel	Orders Corrected	36
Night Shift Workers	Hotel	Reservation Phone Calls Made Orders Corrected	11

In the analysis, the Order Takers, Order Correctors and Night Shift Workers were grouped together.

Appendix A2: Comparison between treatment and control groups

	Treatment	Control	Stan Dev.	p-value
Number	131	118		
Age	24.44	24.35	3.55	0.85
Male	0.47	0.47	0.50	0.99
Secondary technical school	0.46	0.47	0.50	0.80
High school	0.18	0.14	0.36	0.39
Tertiary	0.35	0.36	0.48	0.94
University	0.02	0.03	0.15	0.34
Prior experience (months)	18.96	16.76	25.88	0.50
Tenure (months)	26.14	28.25	21.92	0.45
Married	0.22	0.32	0.44	*0.07*
Children	0.11	0.24	0.38	*0.01*
Age of the child	0.53	0.71	1.92	0.45
Rent apartment	0.25	0.20	0.42	0.44
Cost of commute (Yuan)	7.89	8.34	6.96	0.61
Internet	0.99	1.00	0.06	0.34
Own bedroom	0.97	0.99	0.14	0.22
Base wage (Yuan)	1539.86	1562.80	161.45	0.26
Bonus (Yuan)	1030.90	1092.59	625.33	0.44
Gross wage (Yuan)	2949.73	3003.36	789.63	0.59
Order takers	68	66		0.93
Order placers	36	32		0.72
Order correctors	19	17		0.87
Night shift workers	8	3		0.23

Notes: Treatment includes employees in airfare and hotel department in November 2010 who are both willing and eligible to participate in the WFH program and have an even birth date (2nd, 4th, 6th, etc. day of the month). Control includes employees in airfare and hotel department in November 2010 who are both willing and eligible to participate in the WFH program and have an odd birth date (1st, 3rd, 5th, etc. day of the month).

Table 1. WFH volunteers were more likely to married, have children, have less education and tenure, and face longer commutes

Dependent variable is volunteer to work from home	(1)	(2)	(3)	(4)	(5)	(6)	(7)	Sample mean
Children	0.123**		0.054	0.075	0.081		0.084	0.08
	(0.056)		(0.083)	(0.083)	(0.083)		(0.084)	
Married[+]		0.095**	0.012	0.054	0.052		0.057	0.15
		(0.044)	(0.065)	(0.066)	(0.066)		(0.068)	
Daily commute (minutes[+])			0.062**	0.062**	0.071**		0.072**	80.6
			(0.030)	(0.031)	(0.032)		(0.032)	
Own bedroom			0.095***	0.088**	0.089**		0.089**	0.60
			(0.035)	(0.035)	(0.036)		(0.037)	
Tertiary education and above				-0.080**	-0.088***		-0.086**	0.42
				(0.033)	(0.033)		(0.034)	
Tenure (months[+])				-0.268***	-0.415***		-0.401***	25.0
				(0.080)	(0.110)		(0.117)	
Gross wage (1000 Yuan)					0.048**	-0.019	0.048**	2.86
					(0.024)	(0.017)	(0.024)	
Age							-0.002	23.2
							(0.007)	
Male							0.010	0.32
							(0.036)	
Number of Employees	994	994	994	994	994	994	994	994

Notes: The regressions are all probits at the individual level of the decision to work from home. Marginal effects calculated at the mean are reported. The total sample covers all CTrip employees in their Shanghai airfare and hotel departments. Willingness to participate was based on the initial survey in November 2010. Employees were not told the eligibility rules in advance of the survey (own room, 6+ months tenure, broadband internet connect). ` Robust standard-errors are reported. [+] denotes the coefficients and standard errors have both been multiplied by 100 for scaling purposes.

Table 2: The performance impact of WFH was positive, about 1/3 from more calls per minute and 2/3 from more minutes

Dependent Variable	(1) Overall Performance	(2) Gross Wage	(3) Overall Performance	(4) Phonecalls	(5) Phonecalls	(6) Phonecalls Per Minute	(7) Minutes on the Phone
Period	Pre and during experiment	Pre and during experiment	During experiment	Pre and during experiment	Pre and during experiment	Pre and during experiment	Pre and during experiment
Dependent Normalization	z-score	log	z-score	z-score	log	log	log
Experiment$_t$*Treatment$_i$	0.232***	0.094***		0.248***	0.120***	0.032**	0.088***
	(0.063)	(0.032)		(0.058)	(0.025)	(0.013)	(0.027)
Treatment$_i$			0.184**				
			(0.086)				
Number of Employees	249	249	249	134	134	134	134
Number of Time Periods	85	20	37	85	85	85	85
Individual Fixed Effects	Yes	Yes	No	Yes	Yes	Yes	Yes
Observations	17806	4648	7476	9426	9426	9426	9426

Notes: The regressions are run at the individual by week level (except column (2) which is run at monthly level), with a full set of individual and week fixed effects. Experiment*treatment is the interaction of the period of the experimentation (December 6th 2010 until August 14th 2011) by an individual having an even birthdate (2nd, 4th, 6th, 8th etc day of the month). The pre experiment period refers to January 1st 2010 until November 28th 2010. During the experiment period refers to December 6th, 2010 and ended on August 14th, 2011. Overall performance is the z-score for each employee on their main task. The z-scores are constructed by taking the average of normalized performance measures (normalizing each individual measure to a mean of zero and standard deviation of 1 across the sample). Since all employees have z-scores but not all employees have phonecall counts (because for example they do order booking) the z-scores for overall performance covers a wider group of employees than for phonecall. Minutes on the phone is recorded from the call logs. Three employees have been excluded because they lack pre-experimental data. Standard errors are clustered at the individual level. *** denotes 1% significance, ** 5% significance and * 10% significance.

Table 3: The increase in minutes from WFH came about two thirds from more minutes per day and one third from more days

VARIABLES	(1) Minutes on the Phone	(2) Minutes on the Phone/Days Worked	(3) Days Worked	(4) Minutes on the Phone	(5) Minutes on the Phone/Days Worked	(6) Days Worked
$Experiment_t*Treatment_i$	0.088***	0.063***	0.025**	0.069**	0.049*	0.021
	(0.027)	(0.024)	(0.012)	(0.030)	(0.027)	(0.013)
$Experiment_t*Treatment_i*$ [total commute>120min]$_i$				0.069*	0.055*	0.014
				(0.036)	(0.031)	(0.017)
Number of Employees	134	134	134	134	134	134
Number of Weeks	85	85	85	85	85	85
Observations	9426	9426	9426	9426	9426	9426

Notes: The regressions are run at the individual by week level, with a full set of individual and week fixed effects. Experiment*treatment is the interaction of the period of the experimentation (December 6th 2010 until August 14th 2011) by an individual having an even birthdate (2nd, 4th, 6th, 8th etc day of the month). The pre experiment period refers to January 1st 2010 until November 28th 2010. During the experiment period refers to December 6th, 2010 and ended on August 14th, 2011. In column (4)-(6), Experiment x Treatment is further interacted with a dummy variable indicating whether an employee's total daily commute (to and from work) is longer than 120 minutes (21.3% of employees have a commute longer than 120 minutes). Standard errors are clustered at the individual level. *** denotes 1% significance, ** 5% significance and * 10% significance. Minutes on the phone are recorded from the call logs.

Table 4: The impact of WFH also looked good benchmarked against Nan Tong and non-experimental employees

VARIABLES	(1) Overall Performance (z-score)	(2) Phone calls (z-score)	(3) Overall Performance (z-score)	(4) Phone calls (z-score)
Comparison group	Nan Tong	Nan Tong	Non-experiment	Non-experiment
$Experiment_t*treatment_i$	0.204***	0.283***	0.302***	0.312***
	(0.047)	(0.048)	(0.060)	(0.064)
$Experiment_t*control_i$	-0.026	-0.009	0.066	0.019
	(0.048)	(0.043)	(0.061)	(0.061)
Observations	100420	86984	27823	15261

Notes: Nan-Tong is CTrip's other large call center, located in Nan-Tong, a city about 1 hour drive outside of Shanghai. This call center also had airfare and hotel departments, and calls were allocated across the Shanghai and Nan Tong call centers randomly. The "Eligible non-experimental group" are the individuals that were eligible for the experiment (own room, 6+ months of tenure and broadband) but did not volunteer to participate in the WFH experiment in the two departments in Shanghai. The regressions are run at the individual by week level, with a full set of individual and week fixed effects. Experiment*treatment is the interaction of the period of the experimentation (December 6th 2010 until August 14th 2011) by a Shanghai based individual having an even birthdate (2nd, 4th, 6th, 8th etc day of the month), while Experiment*control is the interaction of the period of the experimentation by a Shanghai based individual having an odd birthdate. All performance measures are z-scores (constructed by taking the average of normalized performance measures, where these are normalizing each individual measure to a mean of zero and standard deviation of 1 across the sample). Standard errors are clustered at the individual level. *** denotes 1% significance, ** 5% significance and * 10% significance.

Table 5: Selection Effects: Employees that choose to WFH after the experiment ended showed performance improvements almost double the experimental impact

	(1)	(2)	(3)	(4)	(5)	(6)
Dependent Variable	Performance	Performance	Performance	Log(Phonecalls)	Log(Phonecalls)	Log(Phonecalls)
Dependent Normalization	z-score	z-score	z-score	Log	log	log
Sample	All	All	Balanced	All	All	Balanced
$\text{Experiment}_t*\text{WFH}_{i,t}$	0.232***	0.221***	0.174***	0.120***	0.125***	0.104***
	(0.063)	(0.049)	(0.057)	(0.025)	(0.035)	(0.041)
$\text{Post-Experiment}_t*\text{WFH}_{i,t}$		0.284***	0.245***		0.220***	0.203***
		(0.082)	(0.089)		(0.059)	(0.066)
Number of Employees	249	249	150	134	134	73
Number of Weeks	85	144	144	85	144	143
Observations	17806	25449	18214	9426	13278	8866

Notes: $\text{WFH}_{i,t}$ here is defined as working-from home at least one day that week. The regressions are run at the individual by week level, with a full set of individual and week fixed effects. The pre-experiment period is January 1st 2010 until November 28th 2010. Experiment*WFH is the interaction of the period of the experimentation (December 6th 2010 until August 14th 2011) with an individual having worked from home at least one day a week by week. Post-experiment*WFH is the interaction of the period after the experimentation from August 14th 2011 until end of September 2012 with an individual having worked from home at least one day a week by week. Balanced panel drops anybody that quits before the end of March 2012. Individually clustered standard errors *** denotes 1% significance, ** 5% significance and * 10% significance.

Table 6: Employee switches after the end of the experiment: those who performed poorly at home tended to return to the office

Dep. Variable is to Switch Status	(1)	(2)	(3)	(4)
Sample: Treatment (home workers)	Home to Office	Home to Office	Home to Office	Home to Office
Performance during the experiment	-0.075		-0.168**	-0.229***
	(0.058)		(0.079)	(0.082)
Performance before the experiment		0.009	0.143	0.214**
		(0.066)	(0.094)	(0.096)
Married				-0.214**
				(0.071)
Live with parents				-0.186*
				(0.101)
Cost of commute				-0.010
				(0.008)
Observations	110	110	110	110
Sample: Control (office workers)	Office to Home	Office to Home	Office to Home	Office to Home
Performance during the experiment	0.057		0.063	0.063
	(0.083)		(0.109)	(0.111)
Performance before the experiment		0.039	-0.011	-0.022
		(0.106)	(0.141)	(0.146)
Married				0.100
				(0.132)
Live with parents				0.056
				(0.138)
Cost of commute				0.011
				(0.011)
Observations	74	74	74	74

Notes: The regressions are all probits at the individual level. Marginal effects calculated at the mean are reported. Pre-experiment performance is the average of individual weekly performance z-score during the pre-experimental period from January 1st 2010 to November 28th 2010. During experiment performance is the average of individual weekly performance z-score during the post-experimental period from December 6th 2010 to August 14th 2011. Sample for returning to the office includes the 110 treatment workers still at CTrip at the end of the experiment in September 2011; 27 petitioned to come back to the office, yielding a 24.5% return rate. The sample for moving home includes the 74 control group employees still in the experiment by September 1st, 2011. Out of 74 control workers, 27 petitioned to work at home, yielding a 36.5% join rate. Robust standard errors: *** denotes 1% significance, ** 5% significance and * 10% significance.

Table 7: Working from home is associated with improved employee self-reported satisfaction and attitude scores

Variables:	(1)	(2)	(3)	(4)	(5)	(6)
	Satisfaction	General Satisfaction	Life Satisfaction	Exhaustion	Positive Attitude	Negative Attitude
Data source:	Satisfaction survey			Emotion Survey		
Experiment *treatment	0.155***	0.072***	0.168***	-0.564***	0.160***	-0.183***
	(0.058)	(0.024)	(0.053)	(0.168)	(0.040)	(0.058)
Announcement*treatment				-0.102	0.080*	-0.095
				(0.167)	(0.042)	(0.058)
Observations	855	855	855	5109	5109	5109

Notes: The satisfaction survey used in columns (1) to (3) was conducted five times throughout the experimental period: once in early November before the randomization took place and four times after the experiment had started. Treatment defined as having an even birthdate. See details of survey questions and methodology in Appendix A3. The emotion survey used in columns (4) to (6) was conducted every week. The first week was conducted in late November 2010, before the experiment began but after the randomization so that individuals had been informed of their status in the treatment or control groups. All the dependent variables are logged values. The regressions are run at the individual level with a full set of time-dummies. Experiment*treatment is the interaction of the treatment group with the period of the experimentation. Announcement*treatment is the interaction with the treatment group with the period of post-announcement but pre-experiment. Standard errors are clustered at the individual level. *** denotes 1% significance, ** 5% significance and * 10% significance.

Table 8. Attrition: Poor performers are more likely to quit, especially in the control group

Dependent variable: Quit	(1)	(2)	(3)	(4)
Sample	Total	Total	Control	Treatment
Treatment$_t$	-0.172***	-0.053		
	(0.055)	(0.067)		
Performance$_{i,t}$		-0.279***	-0.354***	-0.088*
		(0.063)	(0.078)	(0.047)
Performance$_{i,t}$*Treatment$_t$		0.152		
		(0.096)		
Age	-0.031***	-0.027***	-0.024	-0.027***
	(0.010)	(0.010)	(0.018)	(0.010)
Men	0.074	0.017	-0.022	0.056
	(0.056)	(0.059)	(0.098)	(0.065)
Married	-0.078	-0.128	-0.227	-0.047
	(0.097)	(0.103)	(0.209)	(0.100)
Cost of Commute	0.008**	0.008**	0.011	0.005*
	(0.004)	(0.003)	(0.009)	(0.003)
Children	0.231*	0.414**	0.445*	0.365
	(0.145)	(0.180)	(0.252)	(0.251)
Observations	249	249	118	131

Notes: The regressions are all probits at the individual level. Marginal effects calculated at the mean are reported. The dependent variable is whether the employee quit over the experimental period between December 6th 2010 and August 14th 2011. During this period, control group has an attrition rate of 35.0%, compared to 17.0% in the treatment group. Treatment defined as having an even birthdate. Performance is the average of individual weekly performance z-score during the experimental period from December 6th 2010 to August 14th, 2011 (or the week before they quit in the case of quitters). Performance*treatment is the interaction of the performance measure by an individual having an even birthdate. Cost of commute is measured at daily level in Chinese Yuan (note that 1 Yuan is about 0.15 Dollars). Robust standard errors are reported. *** denotes 1% significance, ** 5% significance and * 10% significance.

Table 9. Promotion: WFH seems to have a promotion penalty conditional on performance

Dependent Variable: Promoted to team leader or more advanced function	(1)	(2)	(3)	(4)
Specification	Probit	Probit	Probit	Probit
Treatment	-0.065	-0.087*	-0.085*	-0.091**
	(0.047)	(0.048)	(0.048)	(0.047)
Performance during experiment		0.086***	0.092***	0.150***
		(0.030)	(0.032)	(0.046)
Performance during experiment*treatment				-0.105*
				(0.061)
Men			0.080*	0.082*
			(0.049)	(0.048)
Tenure (months)			0.001	0.001
			(0.001)	(0.001)
Education: high school and below			-0.002	0.003
			(0.061)	(0.062)
Observations	249	249	249	249

Notes: All regressions are with the dependent variable being whether promoted to team leader or more advanced job function during a 22-month period between Dec 6th, 2010 and Sep 30th, 2012. During this period, a total of 40 employees were promoted, resulting in a 15.9% promotion rate. Performance during the experiment is the average of individual weekly performance z-score during the post-experimental period from December 6th 2010 to August 14th, 2011. For probits, marginal effects evaluated at the mean are reported. Performance*treatment is the interaction of the performance measure by an individual having an even birthdate (2nd, 4th, 6th, 8th etc day of the month). Robust standard errors are reported. *** denotes 1% significance, ** 5% significance and * 10% significance.