



LBS Research Online

[E Sherman](#)

Discretionary Remote Working Helps Mothers Without Harming Non-mothers: Evidence from a Field Experiment
Article

This version is available in the LBS Research Online repository: <http://lbsresearch.london.edu/id/eprint/1064/>

[Sherman, E](#)

(2020)

Discretionary Remote Working Helps Mothers Without Harming Non-mothers: Evidence from a Field Experiment.

Management Science, 66 (3). pp. 1351-1374. ISSN 0025-1909

DOI: <https://doi.org/10.1287/mnsc.2018.3237>

INFORMS (Institute for Operations Research and Management Sciences)

<https://pubsonline.informs.org/doi/10.1287/mnsc.20...>

Users may download and/or print one copy of any article(s) in LBS Research Online for purposes of research and/or private study. Further distribution of the material, or use for any commercial gain, is not permitted.

Discretionary Remote Working Helps Mothers Without Harming Non-Mothers: Evidence From a Field Experiment

Eliot L. Sherman¹
London Business School
esherman@london.edu

October 17, 2018

¹ Direct all correspondence to Eliot L. Sherman: esherman@london.edu; 44 20 7000 8914. I wish to thank Lawrence W. Sherman, Tariq Ahmed, Nick Skinner, Lydia Seymour, the Abcam executive leadership team, all study participants, the associate editor, three anonymous reviewers, and audience members at the 2018 London Business School Organisational Behaviour Faculty Lab. I am grateful to the Leadership Institute at London Business School for supporting this research. The usual disclaimer applies.

The organization of family life constitutes a significant obstacle, if not the primary obstacle, to the realization of gender equity in the workplace (Hewlett 2002, Williams 2010, Petersen et al. 2014).² Compared to men, women spend at least twice as much time providing childcare (Sayer et al. 2004, Lamb and Tamis-Lemonda 2004) and, as a result, experience a significant decline in work hours following motherhood (Bertrand et al. 2010, Killewald and Garcia-Manglano 2016). Furthermore, disproportionate childcare responsibilities compel women to endure sleep disruptions at a greater rate (Maume et al. 2009), refuse travel assignments and late meetings more frequently (Maume 2006), spend many more hours each week multitasking (Sullivan 1997, Offer and Schneider 2011), and miss work more often to provide urgent childcare (Maume 2008). Parenthood is also associated with a considerable increase in household labor for women—but not men—which is, in turn, negatively correlated with earnings (Sanchez and Thomson 1997, Noonan 2001).

It is little wonder, then, that women’s professional attainment continues to suffer, despite the fact that by some measures the human capital of women now exceeds that of men in the aggregate.³ Although it remains unclear exactly how much of the attainment gap is attributable to motherhood, this amount appears to be growing, as “ideal worker” norms prescribing continuous availability (Acker 1990, Kelly et al. 2010) collide with “intensive mothering” norms (Hays 1996), which suggest that mothers are fundamentally deficient unless they prioritize “meeting the needs of dependent children above all other activities” (Correll et al. 2007, p. 1306). By one estimate, for example, the portion of the gender wage gap in Denmark that is due to children doubled to 80% between 1980 and 2013 (Kleven et al.

² The conceptualization of family life, both in this paper and in the research on which I draw, presumes heteronormativity. This constitutes an important scope condition, particularly in light of evidence that homosexual couples allocate childcare more equally than do heterosexual couples (McPherson 1993; Chan et al. 1998).

³ For example, according to the U.S. Census women are more likely than men to have a bachelor’s degree (<https://www.census.gov/newsroom/blogs/random-samplings/2015/10/women-now-at-the-head-of-the-class-lead-men-in-college-attainment.html>, accessed October 2018.)

2018).

Cognizant of the constraints that caregiving imposes on their female employees, many firms maintain “flexibility” policies designed to mitigate conflict between work and family life (for a recent review, see Kossek and Lautsch 2018). However, while flexibility policies take many forms—such as compressed workweeks, job sharing, and reduced hours with pro-rated pay—their ultimate effect is to limit policy users’ presence in the workplace. Accordingly, many women avoid using the reduced-hours policies that are available to them out of concern for the career penalties that would result (Bailyn 1993 [2006], Blair-Loy and Wharton 2002, Briscoe and Kellogg 2011). Such concerns are well founded, as women who use reduced-hours policies experience slower wage growth (Glass 2004), fewer promotions (Kalleberg and Reskin 1995), lower quality work assignments (Stone and Hernandez 2013) and difficulty maintaining relationships with mentors (Turco 2010). Consequently, relatively low uptake of reduced-hours policies by the employees who stand to benefit the most from them has proven to be a “remarkably resilient problem” (Williams et al. 2013, p. 209).

The practice of allowing *all* employees to work remotely at their discretion may constitute a promising alternative.⁴ Even a limited amount of remote working could meaningfully reduce the conflict between work and family demands for women, potentially improve their job performance, and—critically—be appealing enough to men as to result in gender-neutral usage patterns. The latter point is particularly important, insofar as the gendered nature of reduced-hours policy usage perpetuates the very stereotypes regarding commitment and competence that continue to hinder women professionally (Epstein et al. 1999, Blair-Loy 2003, Ridgeway and Correll 2004, Turco 2010).

Accordingly, the goal of this paper is to assess whether discretionary remote working,

⁴ I use the term remote working in lieu of “telecommuting,” as no commuting is involved, and “working from home,” as the distinguishing feature is working outside of the office, which may or may not entail working at home.

as an organizational practice, can ameliorate disparities between mothers and non-mothers in a white-collar workplace. I pursue this goal in three primary ways. First, I assess whether mothers, fathers, and childless employees exhibit different propensities to work remotely when given the discretion to do so. Second, I evaluate the effects of working remotely across these employee subgroups with respect to both the work-family interface and self-reported job performance. Third, I gather evidence regarding the potential coordination costs of remote working, as this speaks to the feasibility of implementing the practice on a long-term basis. Taken together, this approach—in terms of both subject matter and methodology—is consistent with recent calls for scholarship that informs the practice of evidence-based management (Rousseau 2006, Rousseau and Gunia 2016).

The field site for this research was Abcam PLC, a mid-sized life sciences company headquartered in Cambridge, England. I evaluated remote working at Abcam via a four-week randomized experiment with a repeated cross-over design. Reliance upon an experimental framework was critical, since employees who acquire remote working privileges in an ad hoc manner are also likely to differ systematically with respect to their human and social capital from employees who do not (Weeden 2005, Davis and Kalleberg 2006, Hornung et al. 2008, Wharton et al. 2008, Golden 2008).

The distinguishing feature of cross-over experiments is that they allow each subject to receive the treatment and to serve as their own control. Specifically, subjects are randomly assigned to receive either the treatment or control initially, and then “cross over” to the other condition after a fixed amount of time. These switches occur more than once in a repeated trial. This design confers three primary advantages compared with a traditional experimental framework—sometimes referred to as a “parallel group trial”—in which subjects exclusively receive either the treatment or control stimulus (Fisher 1935). First, it obviates equity concerns among the subject population regarding the practice of random assignment (Levitt and List

2009). This is not a trivial matter, as such concerns can preclude an otherwise receptive organization from collaborating. Second, it offers greater statistical power than a parallel group trial with an equivalent number of subjects—a particularly beneficial feature, in light of the propensity for under-powered experiments to produce results that fail to replicate (Simmons et al. 2011), as well as the substantial financial cost of experiments conducted in the field (Broockman et al. 2017). Third, and perhaps most importantly, cross-over designs can recover more accurate treatment effects insofar as they remove fixed subject variance from estimate residuals (Greenwald 1976, Hills and Armitage 1979). A comprehensive statistical discussion of this methodology can be found in Jones and Kenward (2015).

I randomly assigned half of the participants to follow the remote working policy in weeks one and three of the experiment and the other half to follow the policy in weeks two and four, surveying all participants after each week. The remote working policy stated that participants should work from home “as much as was sensible, given [their] professional responsibilities,” whereas during control weeks I instructed participants to work from home as much as they normally would. For most, this entailed coming into the office every day, as the base rate of remote working at Abcam was quite low prior to the trial.

To preview my results, with a 92% completion rate and across 748 person-weeks, I find no significant difference during experimental weeks in self-selection into the number of weekly remote working days across men and women, parents and non-parents, or mothers and fathers. I also find no difference across all subgroups in the total number of hours worked in treatment versus control weeks. In terms of outcomes, remote working significantly reduces family-work conflict for mothers—but not for fathers or non-parents—and significantly improves job performance for most subgroups, though the improvements are especially pronounced for mothers. Remote working also improves job satisfaction for all subgroups with the exception of mothers, whose job satisfaction is not discernibly different during treatment weeks. In terms of

potential coordination costs, remote working reduces co-worker helping among non-parents, but does not have a discernible influence on the co-worker helping behavior of parents. Furthermore, this effect appears during the first half of the experiment but not the second, suggesting that it was attenuated once employees gained more experience with the logistics of working remotely. Finally, interdependence of employees' work does not affect the relationship between remote working and job performance for mothers, though there is some evidence that performance benefits are attenuated for non-mothers with highly interdependent roles. Twenty-two semi-structured interviews conducted with subjects that I selected at random (e.g., Chan and Anteby 2016) corroborate and contextualize these conclusions.

My findings contribute to the work-family and labor market inequality literatures by illustrating, in a field setting where external validity was strong (Morgan and Winship 2007 [2014]), a relatively straightforward way in which organizations can serve as “equalizers” by arresting the reproduction of gender-based inequality within their ranks (Ranganathan and Pedulla 2018). Evidence of this nature remains in short supply, despite decades of attention from scholars: As Dobbin and colleagues (2015, p. 1014) note, “studies of the causes of inequality are legion, but studies of remedies are rare.” This study is also somewhat distinctive insofar as the preponderance of field experiments in labor markets involve audits of hiring practices (Pager 2003, Bertrand and Mullainathan 2004, Correll et al. 2007, Tilcsik 2011, Gaddis 2015, Kang et al. 2016, Rivera and Tilcsik 2017, Weiss Harr 2018, Quadlin 2018). While this work is immensely valuable in its own right, the relative paucity of field experiments that extend beyond the hiring interface limits the accumulation of causal evidence regarding other aspects of the employment relationship. Two recent exceptions (Kelly et al. 2014, Bloom et al. 2015) are particularly germane to the present research. Accordingly, below I discuss in some detail how this project builds on and extends their results.

HOW DOES MOTHERHOOD CONSTRAIN WOMEN PROFESSIONALLY?

Recent scholarship linking the organization of family life to the persistence of labor market inequality emphasizes the irreconcilability of two powerful trends (Pedulla and Thebaud 2015). The first trend involves the proliferation of “ideal worker” norms which compel employees to pursue their professional obligations at the expense of all other commitments, domestic or otherwise (Acker 1990, Williams 2001). This ethos is effectively illustrated by Kelly and colleagues (2010, p. 290), who observed workers debating “whether it is legitimate to leave work to care for a sick child, to take vacation, or to recover from a cold.” Where the ideal worker norm is especially strong—generally speaking, in white-collar professions and managerial roles (Cosser 1974; Jacobs and Gerson 2004)—long hours and uninterrupted availability are viewed as proxies for commitment (Epstein et al. 1999). These effort displays are common precursors to attainment because differences in output quality are often difficult to measure (Blair-Loy 2003, Sharone 2004) and because extensive credentialing may render employees relatively comparable otherwise (Goldin 2014). The diffusion of tournament compensation systems, such as “up or out” promotion tracks in academia, law, and management consulting, contributes to the presence of large discontinuities in pay for workers whose relative rank exceeds that of their colleagues by even a small margin (Connelly et al. 2014). This further compels employees to engage in “overwork,” or workweeks that exceed 50 hours (Cha 2010). Overwork is especially prevalent in the United States, where dual-earner couples log more work hours than their counterparts almost anywhere else in the world (Meyers and Gornick 2005), reflecting a ten percent increase during the last three decades (Mishel 2013).

Ideal worker norms have proliferated in concert with a second trend: The ascendance of “intensive mothering” as a cultural imperative that exacerbates long-standing inequities in the division of domestic labor (Hays 1996). Intensive mothering departs from gender-

essentialist beliefs regarding the superiority of female caregiving—beliefs that remain stubbornly prevalent, even among those who otherwise endorse gender egalitarianism (Cotter et al. 2011)—to suggest that mothers are fundamentally deficient unless they prioritize “meeting the needs of dependent children above all other activities” (Correll et al 2007, p. 1306).

The extent to which intensive mothering transcends boundaries of class, race, and income remains subject to debate. Recent scholarship suggests that the central imperative of perpetual self-sacrifice affects women irrespective of class, as Hays (1996) originally suggested. That is, low-income mothers “embrace and perform intensive mothering in the absence of larger social supports for their children’s upbringing and at a cost to their own emotional and physical well-being” (Elliott et al. 2015, p. 351, see also Edin and Kefalas 2011). Other research suggests that some of the most prominent manifestations of intensive mothering, such as time-consuming intervention in children’s education, are primarily hallmarks of the middle class (Lareau 2000, Vincent and Ball 2007).

Without adjudicating between these perspectives, it is sufficient to observe that intensive mothering, as an ideology, motivates an array of behaviors that are, to a greater or lesser degree, effectively incompatible with white-collar work as it is presently organized (Nelson 2010). Prominent examples include long-duration breastfeeding (Gatrell 2007, Rippeyoung and Noonan 2012) and co-sleeping (Sears and Sears 1992)—both of which tend to excuse men from routine sleep disruption (Maume et al. 2009, 2010)—as well as a time-intensive approach to meal consumption, which emphasizes home preparation and the use of organic materials for the express purpose of mitigating health risks (Moisio et al. 2004, Afflerback et al. 2013).⁵ Likewise, with respect to older children, this ideology compels

⁵ The link between home-prepared organic food and mitigated health risk remains dubious, particularly as there is no scientific consensus regarding the definition of organic (Lyons et al. 2001).

mothers to engage with a wide range of children's extracurricular activities, in addition to the routine medical appointments and vehicular transport that permeate familial life (Lareau 2002, Stone 2008). At least in part as a result of this, the amount of time that mothers presently spend with their children is equivalent to the 1960s, when working full-time was the exception instead of the norm (Bianchi et al. 2006).

Professional Responses to Motherhood

The collision of ideal worker and intensive mothering norms evokes guilt and emotional anguish in women (Simon 1995, Glavin et al. 2011). It also makes attainment far more difficult for working mothers than for men or childless women, particularly in light of men's inability or unwillingness to assume greater childcare responsibilities. This dynamic is implicated in two common responses to childbirth exhibited by professional women. The first is to exit the labor market entirely, at least temporarily. For example, Cha (2010, p. 324) found that "the odds of quitting are 112 percent higher for mothers in professional occupations when their husbands work long hours," while 60% of the women interviewed by Stone (2008) cited their husband's unavailability for childcare in their decision to leave their professional careers. Although many women who quit the labor market after childbirth later return, even brief interruptions in employment can negatively affect their attainment for many years to come (England et al. 2016, Wilde et al. 2010). Consistent with this, a large-scale audit study recently found that mothers who "opted out" of working for familial reasons were less likely to be called back regarding subsequent job applications than both continuously employed women and women who had lost their previous job due to non-familial reasons (Weissarr 2018). Evidence that mothers can regain their lost earnings power by their 40s and 50s (Kahn et al. 2014) may, meanwhile, provide little comfort.

A second common response to motherhood for professional women, which is also

conditioned at least in part on men's relative unavailability for childcare, is to work in a part-time or reduced-hours capacity (Becker and Moen 1999, Stone 2008). This is generally suboptimal for several reasons. First, there is evidence that women participating in reduced-hours policies are often compelled to work more than their designated hours (Blair-Loy 2003, Webber and Williams 2008), with the balance effectively constituting free labor. Second, in many cases the relationship between part-time hours worked and earnings is non-linear (Goldin 2014), such that workers earn less—sometimes substantially less (Bardasi and Gornick 2008)—than their pro-rated full-time equivalent pay. Third, reduced-hours participation is, by definition, incompatible with “overwork.” This is problematic because the financial returns to overwork have increased dramatically during the last forty years, to the extent that men's disproportionate propensity for overwork effectively offset the wage gains made by women as a result of their comparatively greater educational achievements (Cha and Weeden 2014).

Perhaps most detrimental, however, is the fact that even a short reduced-hours stint can substantively alter the way a white-collar worker is viewed by her co-workers and supervisors. At issue is the belief that any deviation from full-time employment contravenes ideal worker norms (Acker 1990), adherence to which constitutes an informal but unyielding prerequisite for career progression. Specifically, insofar as role commitment—manifested via continuous availability—is the core ideal worker tenet (Correll and Benard 2006), part-time workers are denigrated as “time deviants” (Epstein et al. 1999) whose behavior renders them unworthy of further advancement (Blair-Loy and Cech 2014). For example, reduced-hours workers are often transferred to less impactful assignments and face the prospect of mentors withdrawing essential support (Turco 2010, Stone and Hernandez 2013). Likewise, in one finance firm female reduced-hours workers were evaluated as lower-performing than co-workers who did not use the policies (Wharton et al. 2008), while women who used some portion of their sick leave when their children fell ill had significantly lower earnings compared with otherwise

equivalent managers who did not (Blair-Loy and Wharton, 2004). It is, in large part, the stigmatizing effect of reduced-hours policy use that motivates many professional women to quit rather than work part time (Stone 2008, Williams et al. 2013).

The evaluative discount that reduced-hours workers suffer is sustained by the fact that usage is highly gendered (Ridgeway and Correll 2004), to the extent that—in some white-collar workplaces—women’s ideal worker commitment is perennially suspect solely by virtue of their potential for motherhood, and with it, policy use (Turco 2010, Rivera and Tilcsik 2017). Even in less extreme work settings, reduced-hours policy use is sufficiently confirmatory as to trigger the host of negative stereotypes that subsequently hinder attainment (Benard et al. 2007). It remains to be seen, however, whether such stereotypes would obtain from the use of a policy that was more gender neutral in its uptake—and that may, as a result, be more compatible with the ideal worker image. As Thebaud and Pedulla (2016, p. 611) note, “Men’s responses to supportive work–family policies depend largely on their perceptions of what they believe their male peers want, and by extension, what kind of behavior they would hold them accountable to.”

Remote Working as an Organizational Intervention

Evidence regarding the consequences of remote working remains decidedly mixed. On one hand, studies often suggest that increasing employees’ schedule control reduces the conflict between work and family demands, among other beneficial outcomes (Galinsky et al. 1996, Voydanoff 2004, 2007, Grzywacz et al. 2008, Gareis and Barnett 2002). On the other hand, research also suggests that the benefits of greater schedule control may be a mirage, or at least limited—and that, instead, such practices constitute an insidious way for white-collar workplaces to wring a greater number of work hours from their employees than would otherwise be possible (Briscoe 2007, Blair-Loy 2009).

Findings remain mixed for several reasons. First, studies often conflate different types of policy use, which complicates the interpretation of their results. As Glass and Finley (2002, p. 323) note in their review, “A major measurement issue that emerged was the tendency to group family-responsive policies together in the analysis, obscuring which particular policy or combination of policies was affecting the reported outcomes. This occurred most often with regard to flexible work arrangements.” For example, there is correlational evidence that remote working—as distinct from other forms of flexibility—can provide psychological benefits to women with children in particular (Kossek et al. 2006) without constraining earnings growth for either sex (Glass and Noonan 2016).

The second reason that scholarship continues to produce mixed results may be attributable to the relatively high incidence of correlational studies. This constitutes a non-trivial methodological limitation, insofar as the antecedents and consequences of remote working vary systematically as a function of employees’ human and social capital (Weeden 2005, Davis and Kalleberg 2006, Hornung et al. 2008, Wharton et al. 2008, Golden 2008). Two experimental studies constitute exceptions that are particularly germane; accordingly, I review them here in some depth. First, Kelly and colleagues (2014) conducted a group-randomized trial called STAR with employees in the IT department of a U.S.-based Fortune 500 company. Second, Bloom and colleagues (2015) conducted an experiment with call-center workers in China.

The STAR experiment was designed, in part, to reduce work-family conflict by increasing employees’ schedule control and supervisors’ support for employees’ family and personal matters. Treatment entailed a four-hour training session for supervisors, encouraging them to demonstrate support for employees’ personal lives, and an eight-hour training session for employees and supervisors, in which participants discussed “new ways of working to increase employees’ control over their work time and demonstrate greater support for others’

personal obligations” (Kelly et al. 2014, p. 490). A six-month follow-up survey of 717 participants indicated that STAR significantly reduced family-work conflict, increased employees’ schedule control, and increased the amount of hours per week that participants worked from home, among other outcomes.

While the STAR experiment is comprehensive and compelling, the treatment was—consistent with its goal of changing the firm’s cultural schemas (Sewell 1992)—multi-faceted and complex. That is, while experimental groups were effectively exposed to the same treatment, the manner in which they capitalized on that treatment varied, insofar as groups decided among themselves how best to act on what they had learned. As a result, it is not clear to what extent remote working specifically can be implicated in the reduction of family-work conflict—as opposed to schedule control more generally or other aspects of the treatment, such as increases in emotional support from supervisors and co-workers.⁶ In addition, the question of employee performance was beyond the scope of the investigation reported in both the initial trial and its subsequent follow-up (Moen et al. 2016).

Bloom and colleagues (2015) examined the impact of remote working specifically with respect to employee performance, finding generally positive effects in the context of call center work. Accordingly, it is useful to highlight several points of divergence between their experiment and the present study. First, Bloom and colleagues conducted their experiment in China, which remains in many ways distinct from Western professional culture (e.g., Otis 2008). Second, Bloom and colleagues required subjects in the treatment condition to work from home for four days per week for the entirety of the trial. This represents an extreme version of remote working: For example, one review characterized remote working for more than 2.5 days per week as “high-intensity” (Gajendran and Harrison 2007). By comparison, my design elicits

⁶ For example, the authors measured schedule control via an eight-item scale, of which only one item pertained to remote working.

workers' revealed preferences with respect to the optimal number of remote working days, which was half of the amount required by Bloom and colleagues. Third, while Bloom and colleagues' experiment measured employee performance very effectively, they did not investigate the effects of remote working on the work-family interface. Finally, as Bloom and colleagues note, "The job of a call center employee is particularly suitable for telecommuting. It requires neither teamwork nor in-person face time," conditions which are "far from universal" (Bloom et al. 2015, p. 213). As a result, it is still unclear how remote working affects employees whose professional responsibilities are less portable, which constitutes a sizeable portion of white-collar work. Assessing the potential coordination costs of remote working for these employees is crucial, however, insofar as it speaks to the organizational feasibility of maintaining such a policy over the long term.

METHODS

Randomized cross-over design. I investigated the uptake and consequences of remote working using a randomized field experiment with a repeated cross-over design. Cross-over experiments are commonly used in medical trials (e.g., Hills and Armitage 1979, Martins et al. 2018), as they economize statistical power and thereby reduce the cost of subject recruitment. But they are employed much more infrequently in the social sciences—so much so, in fact, that "experiment" is often synonymous with parallel group trial, in which subjects are randomly exposed to either a treatment or control stimulus, with analysis entailing a comparison of means (Fisher 1935).⁷ By contrast, in a cross-over study subjects are randomly assigned to a sequence

⁷ The defining feature of an experiment is not random assignment to parallel groups but, rather, the presence of an exogenously-induced manipulation. As Jackson and Cox (2013, p. 28) note, "Broadly, the investigation is an experiment if the investigator controls the allocation of treatments to the individuals in the study and the other main features of the work, whereas it is observational if, in particular, the allocation of treatments has already been determined by some process outside the investigator's control and detailed knowledge."

of treatments that occur over a set number of periods. In a balanced design, which I employ, subjects are allocated an equivalent number of sequences, periods, and treatments. Specifically, I randomized subjects to either an ABAB or BABA sequence, where A indicates a treatment week and B indicates a control week.⁸ For treatment weeks, I instructed participants to “relax the assumption that work must necessarily be done in the office and, instead, work remotely for as much as is sensible given your professional responsibilities.” For control weeks, I instructed participants to proceed as usual, that is, to “work in the office as much as you would normally do.” I informed participants of their assigned sequence in the week before the experiment began.

Cox (1958) and Greenwald (1976) appraise the strengths and weaknesses of this type of design—sometimes referred to as “within-subjects” by psychologists—in comparison to a traditional parallel group trial. In terms of threats to internal validity, the potential for a “carry over” effect, which “occurs when the effect of one treatment persists in some fashion at the time of measurement of the effect of another” (Greenwald, 1976, p. 318), constitutes the primary difference. At issue in medical trials is the potential for trace amounts of the drug under study to remain with subjects who received it during the first period when measures are taken from them during the subsequent placebo period. With respect to remote working, the concern would be that—for example—the psychological benefits that accrued to subjects who received the treatment in week one effectively “carried over” to their subsequent control week, to the extent that they are no longer comparable to subjects randomized to the alternative sequence.

In order to address this concern, cross-over experiments usually separate treatments temporally by inserting a “wash-out period” between them (e.g., Farr et al. 2017). In this case,

⁸ I used STATA to produce simple random assignment of the entire sample, generating random numbers for all participants, ranking the numbers in descending order, and then splitting them in half to separate ABAB participants from BABA ones.

the weekend served as my washout period. No problems arise if it can be reasonably assumed that this is sufficient to preclude the initial treatment from affecting subsequent ones (Cox 1958), while for some investigations a wash-out period may not be necessary (Martens et al. 2018). A useful feature of cross-over experiments, however, is the ability to test for carry-over effects directly, in order to determine whether the wash-out period was in fact sufficient. Below I describe the steps that I took in order to do so: In short, I found no evidence of a treatment-by-sequence interaction effect. This indicates that the wash-out period was effective.

Subjects and recruitment. My research site was Abcam PLC, a mid-sized life sciences firm headquartered in Cambridge, England. The company was founded in 1998 with the goal of facilitating the worldwide sale of antibodies and related products, such as peptides and proteins, via the Internet. Its customer base is primarily research scientists. In 2016, Abcam earned approximately £50 million on revenues of £170 million. It has been listed on the United Kingdom's Financial Times Stock Exchange (FTSE) since 2005.

At the time of the experiment Abcam employed approximately 1,000 people worldwide, of whom 434 were located in Cambridge. To assist with recruiting subjects in the Cambridge office, I asked line managers from the company's different departments to disseminate an e-mail that I wrote informing eligible employees about the experiment and endorsing their participation (see **Appendix A**). The only employees who were deemed ineligible were those whose responsibilities exclusively required their physical presence onsite. In effect, this meant certain logistics workers involved in handling, packaging, and storing materials.

In order to register for the experiment, participants filled out a baseline survey in which they reported demographic information and responded to the scale items that constituted my independent and dependent measures. 203 employees, or 45% of the workforce in Cambridge,

signed up.⁹ In order to encourage participation—and, importantly, to avoid subject attrition during the experiment—I offered each participant a £50 gift voucher to John Lewis, a popular retailer, conditional on the completion of all four weekly surveys subsequent to the baseline survey. Ensuring a high level of completion is crucial for field experiments, because the systematic attrition of subjects can complicate the analysis and interpretation of experimental results (Levitt and List 2009, Baldassarri and Abascal 2017). Of the 203 employees who signed up initially, 187—or 92%—completed all four weekly surveys, and therefore serve as the basis for the analysis I report below.¹⁰ I attribute this completion rate, which is comparable to the rate reported by Kelly and colleagues (2014), to the financial inducement I provided and the brevity of the weekly surveys that I assigned, each of which took approximately four minutes to fill out.

T-tests indicated that participants and non-participants were not statistically distinguishable by gender ($p = 0.98$). Participants did earn less on average than non-participants (£44,730 versus £60,980).¹¹ However, this difference is attributable to the fact that the highest earners in the firm, such as the CEO and the CFO, did not participate in the experiment. Conditioning the comparison on those earning less than £200,000 per year, for example—which excludes just the twelve highest earners in the firm—renders participant and non-participant salaries statistically indistinguishable. Thus the external validity of the volunteer

⁹ I assigned 102 participants to the ABAB sequence and 101 participants to the BABA sequence. In my analytic sample, 93 participants received the ABAB sequence and 94 received the BABA sequence.

¹⁰ An additional 25 employees expressed interest in joining the experiment during the first week. In response, I assigned them to the BABA sequence, in order to avoid circumscribing their potential number of remote working days during their first exposure to the treatment. Of these 25 employees, 24 completed all four weekly surveys. I conducted the analysis that I describe in this section with and without their data included. The inclusion of their responses did not meaningfully alter the pattern of results that I observed; accordingly, I excluded their responses from all the models that I report, and restricted the data to those subjects who were randomized to their respective sequences.

¹¹ The salary comparisons are missing data from 13 study participants, for whom earnings information was not available. Overall the average annual salary at the firm was £54,428, with a standard deviation of £50,656.

sample, with respect to the company overall as well as similar companies, should be relatively strong.

The experiment ran for four weeks, beginning on Monday, September 25, 2017, with outcomes measured via online surveys at the end of each week. Surveys went live at noon on Friday and closed at 9 AM on the following Monday. In terms of race and nationality, subjects were predominately white and British. The average participant age was 34 years; 52% were female; 37% had at least one child in their household; 44% lived with a spouse; and 39% could reasonably be classified as managers, in that they supervised at least one person who reported directly to them. At baseline, the average number of weekly working hours was 38. See **Table 1** for a full list of descriptive statistics.

The average number of weekly remote working days prior to the experiment was 0.33, but this distribution was highly skewed: 67% of employees reported that they never worked from home, 11% reported working from home a half day per week, 18% reported working from home one day per week, and 4% reported working from home for more than one day per week. Importantly, however, subjects did not exhibit significant variation in remote working propensity by subgroup prior to the experiment (see **Table 2**).

******* Insert Tables 1 and 2 About Here *******

Measures.¹² My assumption was that longer surveys would result in a higher rate of subject attrition. Accordingly, I only included measures that I deemed absolutely essential to the project. Further, in order to increase comparability with prior research, wherever possible I included the same scale items that Kelly and colleagues (2014) used. Accordingly, with respect to *family-to-work conflict*, I utilized the five-item scale developed by Netermeyer and

¹² In addition to what I report in this section, I collected—but did not use—the following two measures. First, in the baseline survey that subjects took prior to the experiment, I included one question that assessed respondents' level of elder care responsibilities. These were minimal overall. Second, at Abcam's request, in the baseline survey and the weekly surveys I included a question which measured how likely respondents were to recommend working at Abcam to a friend.

colleagues (1996). The scale includes items such as, “Things I wanted to do at work didn’t get done because of the demands of my family or personal life.” I focused on the *family-to-work* direction because remote working is presumably beneficial insofar as it limits the demands of home life—such as a sick child—from precluding the completion of work. Brevity concerns precluded me from measuring *work-to-family* conflict as well, though this direction also appeared less germane to remote working specifically.

I measured *job performance* via the three-item scale developed by Welbourne and colleagues (1998), which includes items such as, “How would you evaluate your job performance with respect to the quality of your work output?” I measured *job satisfaction* using the three-item scale developed by Camman and colleagues (1983), which includes items such as, “In general, I like working at my job.” In order to evaluate whether remote working created coordination costs, I measured co-worker *helping* behavior via the four-item scale developed by Dekas and colleagues (2013), which includes items such as “I helped others who had heavy workloads.” I also measured *job interdependence*—at baseline only—via the five-item scale developed by Van der Vegt and colleagues (2001), which includes items such as “I have to obtain information and advice from my colleagues in order to complete my work.”

I measured all scale items from one to five, anchoring on “strongly disagree” and “strongly agree,” respectively. See **Appendix 2** for a complete list of these scale items. In terms of non-scale measures, I assessed weekly *remote working days* from zero to five in half-day increments; weekly *work hours*; weekly *flex hours*, that is, hours worked outside of 8:00 AM to 6:00 PM; and weekly *leave days*, that is, sickness or vacation days taken.

Manipulation check. During the experiment subjects reported working an average of 2.14 days remotely during treatment weeks versus 0.49 days remotely during their control weeks, a more than three-fold relative increase. A T-test indicates that this constituted a significant difference ($p = 0.000$), meaning that the manipulation was successful.

RESULTS

I first conducted a series of T-tests in order to determine whether the uptake of remote working days exhibited significant variation with respect to gender or other demographic differences (see **Table 3**). I observed no clearly significant difference at the 0.05 level in the number of remote working days taken during treatment weeks by gender ($p = 0.06$), parenthood ($p = 0.07$), fatherhood ($p = 0.24$), or motherhood ($p = 0.28$). Furthermore, while it is not a statistically significant comparison, mothers actually took slightly fewer remote working days than non-mothers (1.95 versus 2.18) and parents took slightly fewer remote working days than non-parents (1.95 versus 2.25). Women did, however, report taking more remote working days overall than men (2.28 versus 1.98), though again this difference was not statistically significant. The key point, though, is that the magnitude of the policy effect was not meaningfully different: Abcam employees elected to work remotely about two days per week during their treatment weeks, compared with a half day during control weeks. **Figure 1** shows the uptake of remote working days during treatment and control weeks for each employee subgroup.

***** **Insert Table 3 and Figure 1 About Here** *****

Given the frequency with which children's health and related circumstances require unanticipated childcare during the workday, and the extent to which this is disproportionately addressed by mothers (Maume 2008, Kelly et al. 2010), I examined whether subjects took fewer days of sick or holiday leave during remote working weeks. T-tests indicated that remote working weeks did not influence the number of leave days taken with respect to the full sample ($p = 0.36$). T-tests also indicated that subjects did not log significantly different *work hours* in total during remote working weeks versus control weeks (36.47 versus 36.18, $p = 0.66$). This is consistent with Kelly and colleagues (2014), who did not observe a "work intensification"

effect as a result of increased schedule control (cf. Schieman 2013). Subjects did report working more *flex hours*—that is, hours outside of 8:00 AM to 6:00 PM—during remote working weeks versus control weeks (3.04 versus 1.37, $p = 0.000$).

In terms of *family-work conflict*, *job performance*, *job satisfaction*, and *helping*, I estimated the within-person treatment effect of being in a remote working week versus a control week. I selected this modeling approach because it evaluates the effects of a change in remote working policy, which both addresses my research question directly and is what I randomized. I did not randomize the number of remote working days that subjects took during their treatment weeks—but, as I note below, estimating my dependent variables as a function of this recovers a similar pattern of results.

Accordingly, I regressed my outcomes of interest on a binary indicator of treatment versus control week for each participant with fixed effects for individual and week included.¹³ In each case, I estimated the effect for the full sample at first and then solely for the relevant employee subgroups.¹⁴ The results of these regressions, and the corresponding effect sizes, are shown in **Tables 4 to 7** below. Following Kelly and colleagues (2014) I characterize effect sizes using Cohen's d (1988), which is a between-subjects measure obtained by dividing the relevant coefficient estimate by its standard deviation. Cohen (1988) classified effect sizes of 0.2 as small, 0.5 as moderate, and 0.8 as large, but also cautioned that these are not rigid benchmarks and must instead be interpreted in light of the research question, study context, and prior findings (see also Prentice and Miller 1992).

******* Insert Table 4 and Figures 2 and 3 About Here *******

¹³ One distinction between this approach and the approach employed by Kelly and colleagues (2014) is that I do not incorporate the baseline survey measures into my analysis, as they do by measuring the effect of the intervention as the change from baseline. Rather, I utilize the cross-over design to compare the effects of within-person changes from treatment to control weeks during the four weeks that the experiment was running.

¹⁴ Estimating the models reported in **Tables 4 – 7** without the inclusion of weekly fixed effects does not substantively alter the pattern of results that I observe.

Table 4 shows the effect of remote working on *family-to-work conflict*. The effect is not statistically significant across the full sample, but the subgroup differences are pronounced. In particular, I observe a small and marginally significant ($p = 0.051$) reduction in family-work conflict for parents. However, this effect is driven by mothers—for whom the effect is medium by Cohen’s (1988) classification and clearly significant—and is not statistically significant for fathers. Perhaps unsurprisingly, non-parents do not exhibit any change in *family-to-work conflict* as a result of remote working. See **Figure 2** for a between-subjects comparison of means across subgroups.

Another way to think about the impact of this policy on mothers is to compare the mean levels of *family-to-work* conflict reported by mothers during treatment weeks to the mean levels reported by non-mothers of both sexes during the control weeks. These means are, respectively, 1.58 and 1.55. In other words, the effect of the policy is to render the amount of family-work conflict experienced by mothers effectively equivalent to the levels exhibited by all other employees. By comparison, mothers’ mean family-work conflict during control weeks in the experiment was 2.08, which is the highest of any subgroup (see **Figure 3** for an illustration). This is particularly meaningful in light of a separate within-subjects regression, which I report here but do not include in my main analysis, that indicated a negative relationship between *family-work conflict* and *job performance* across the full sample ($\beta = -0.18, p = 0.000$).

***** **Insert Table 5 and Figure 4 About Here** *****

Table 5 shows the effect of remote working on *job performance*. There is a small but positive effect across the full sample, as well as within the subsamples of men, women, and parents. Notably, however, the effect is largest—medium, per Cohen’s (1988) classification, and clearly significant—for mothers. No significant effect obtains in the subsample of fathers, indicating that the effect on parents is driven by mothers. See **Figure 4** for a between-subjects comparison of means across subgroups.

***** Insert Table 6 and Figure 5 About Here *****

Table 6 shows the effect of remote working on *job satisfaction*. These results are somewhat surprising, in that women generally and mothers specifically do not report a positive and at least marginally significant effect, despite seeming to benefit the most from the policy in terms of reduced *family-to-work conflict* and increased *job performance*. Instead, remote working did not discernibly influence the job satisfaction of women or mothers. The fact that men reported significantly greater *job satisfaction* from remote working while women did not is inconsistent with Moen and colleagues (2016), who did not observe a gender difference in *job satisfaction* as a result of the STAR intervention. Interestingly, remote working has a small but clearly significant impact on the *job satisfaction* of non-parents, a point to which I return in the discussion section. See **Figure 5** for a between-subjects comparison of means across subgroups.

***** Insert Table 7 and Figure 6 About Here *****

Lastly, I assessed the potential co-worker coordination costs of remote working in two different ways. The first is shown in **Table 7**. Remote working has a small but significant negative effect on *co-worker helping* for both men and women. Notably, however, this effect is driven by non-parents, as the effect is not statistically significant for parents, mothers, or fathers. See **Figure 6** for a between-subjects comparison of means across subgroups.

In addition, I investigated whether the effect of remote working on *co-worker helping* obtained during both the first and second halves of the experiment. By regressing co-worker helping on a series of dummy variables for each week with person fixed effects included and week one as the reference category, I found that helping behavior in week two was not significantly different from week one. Yet helping behavior in weeks three and four were significantly higher than week one. Consistent with this, I estimated a model that excluded weeks one and two; that model indicated that there was no statistically significant effect of

remote working on helping behavior ($p = 0.493$). This pattern of results suggests that the negative effect of remote working on *co-working helping* receded once employees gained more experience with the logistics of working remotely, an insight that was corroborated with my interview subjects as noted below.

The second way I assessed potential coordination costs was by re-estimating the models from **Tables 4** and **5** conditional on each subject's self-reported level of job interdependence with co-workers. Specifically, I re-estimated the models conditional on being greater than or equal to, or below, the median of 3.2. With respect to *family-to-work conflict*, a concern might be that the benefits of remote working do not accrue to mothers whose responsibilities entail an above-average level of interdependence with others. This was not the case, however, as the effect was comparably significant for mothers above and below the median ($p = 0.014$ and $p = 0.000$, respectively) and essentially equivalent in size ($\beta = -0.52$ and $\beta = -0.55$, respectively). Likewise, in terms of *job performance*, the effect remained significant for mothers above and below the interdependence median ($p = 0.002$ and $p = 0.013$, respectively) and was essentially identical in size ($\beta = 0.36$ and $\beta = 0.35$, respectively). With that being said, the *job performance* benefits of remote working were attenuated for the full sample of workers whose interdependence was above the median ($\beta = 0.08$, $p = 0.09$), though there was no difference for *family-work conflict*.

Robustness Checks

Carry-over effects. I tested for the presence of carry-over effects by investigating whether selection into, and the effects of, remote working exhibited variation by sequence—that is, the order in which I randomly assigned subjects to treatment versus control weeks.¹⁵

¹⁵ This is sometimes referred to as a “treatment-by-period” interaction for cross-over experiments that are not repeated.

With respect to the former, I estimated the number of remote working days taken as a function of remote working week, sequence, and an interaction between the two, as well as individual fixed effects. The interaction effect was not statistically significant ($p = 0.70$), indicating that the effect of remote working condition on the number of remote working days taken did not differ based on the sequence to which I assigned subjects.

Likewise, in order to test whether carry-over influenced the effect of remote working on *family-work conflict*, *job performance*, *job satisfaction*, and *co-worker helping*, I regressed each respective measure on a binary indicator of remote working week, a binary indicator of sequence, and an interaction between the two, as well as individual fixed effects. In all cases the interaction effect was not significant, returning p -values of 0.75, 0.58, 0.87, and 0.11, respectively. In sum, the lack of a significant interaction effect demonstrates that the effect of the treatment was not different for subjects who received one sequence versus the other, indicating an absence of carry-over.

Dose-response relationship. As noted above, I estimated the effect of being in a treatment versus control week because this is what I randomized. However, as a robustness check, I also estimated variation in *family-work conflict*, *job performance*, *job satisfaction*, and *helping* behavior as a linear function of the number of remote working days taken. The results of these regressions, which are consistent with my main analysis, are shown in **Appendices 3 to 6**.

Possible Hawthorne effects. I also considered whether my results were influenced by a “Hawthorne effect.” Broadly speaking, the Hawthorne effect refers to the potential for worker productivity—or any other outcome under study—to be influenced by subjects’ knowledge that they are being observed, independent of any exogenous manipulation introduced by the researcher. The term is derived from a series of experiments that the National Research Council conducted in Western Electric’s Hawthorne Plant, in Cicero, Illinois, beginning in 1924 (e.g.,

Roethlisberger and Dickson 1939 [2003]). The initial and most well-known experiments investigated whether better lighting could increase worker productivity, while later experiments altered the room size, introduced fans, and relaxed the regulations around talking, among other interventions. Subsequent summaries of these experiments (e.g., Ruch and Zimbardo 1971, Blalock and Blalock 1982) incorrectly asserted that worker productivity increased with every exogenously-induced manipulation, suggesting that subjects' awareness of participating in an experiment constituted a substantial confound. In contrast, re-analyses of the original data uncovered no evidence of a Hawthorne effect at the Hawthorne plant (Jones 1992, Leavitt and List 2011). As Zizzo (2010: 79) states, "The comparative weakness and ambiguity of the evidence implies that, notwithstanding its enduring textbook appeal, it can hardly be used to argue for [experimenter demand effects] being an all pervasive confound."

The issue of subject awareness nevertheless remains a concern for field experiments. The most straightforward way to address this empirically is to include a separate control group as a "placebo," and provide them with a purposefully inert treatment in order to determine whether the mere act of being studied constitutes a confound (e.g., Nyhan and Reifler 2015). The sample size demands of this approach, however, are considerable, and unfortunately could not be met in the Abcam experiment. An alternative, suggested by Orne (1973), involves "postexperimental inquiries" designed to elicit any suspicions that subjects may have maintained during the course of the experiment. I pursued this approach via 22 semi-structured interviews, as I describe below. None of these conversations indicated that participants had experienced "special attention" as a result of their participation (Jones 1992) or been aware of and responsive to my hypotheses *ex ante*.

Contextualizing Quantitative Results with Qualitative Data

In order to corroborate and contextualize my quantitative results, I conducted twenty-two semi-

structured interviews with subject participants I selected at random (Chan and Anteby 2016). In each interview, I asked subjects to describe their professional responsibilities at Abcam, as well as their positive and negative experiences during the experiment. I probed further with respect to family-work conflict, job performance and coordination costs specifically insofar as this information did not emerge organically from subjects' initial responses.

Family-work conflict. The parents that I interviewed were unanimous in their belief that remote working made it easier to manage work and family life. Perhaps unsurprisingly, the views of mothers were especially pronounced in this regard. Amelia, a 33-year-old accounting manager, noted "It had such a positive impact. You know, my children are six and three, so evenings are always a stressful time, and when I'm not leaving the office till six PM, not getting home until seven PM most nights, it puts you behind before you've even started."¹⁶

Amanda, a 38-year-old in the human resources department, agreed:

Some days are easier for me than others. But, say like on a Tuesday, I have got to leave the house at 7:30 AM, drop one child off in one place, the other child off in the next, and I don't actually get into work until 9:00 AM, so I have already done an hour and a half just driving around. I am generally, if the traffic is bad, rushing in, so I'm feeling on edge already. And, then I have got all of that at the back end of the day as well. Whereas, when I was at home, it was closer, instead of taking me an hour and half to sort the kids out, I could do that in half an hour ... So, I think I felt a lot calmer, and as a result I probably approached my work in a much calmer way, as opposed to this sort of, frantic juggle.

Fathers also articulated work-life benefits, though they tended to emphasize the fact that remote working increased their presence and engagement with family members as opposed to allowing them to contribute more logistically. As Oliver, a 34-year-old in a supply management role, described it:

Family-wise, [remote working was] great. Yeah, I'm essentially here for an extra hour because I'm obviously not commuting. So that's good. It's just easier. Work finishes and I'm already home. And then [my family gets] a better version of me. Usually the first thing they see of me is me tired, off my bike and wanting to have five minutes to get myself sorted out. So yeah, they probably get a more awake, more calm me, which is good.

¹⁶ I use pseudonyms throughout in lieu of subjects' real names.

While non-parents reported no reduction in *family-work conflict*, many did point out that the ability to work remotely did help them manage the interface between work and other aspects of their lives. As Jason, a 27-year-old in the information technology group, stated: “It’s not the same as having kids to take care of or take to school or whatever, but everybody has stuff that needs to get done every day, so it just helps with that ... And, you know, being able to get a bit more sleep. I think I noticed a massive benefit from doing that, absolutely.” Some respondents also linked the additional sleep they were able to get during remote working weeks to increased job performance, as described below.

Job performance. Almost all interview subjects reported experiencing productivity gains from their remote working weeks, whereas the few that did not viewed the effect of remote working as purely neutral. Subjects articulated two primary mechanisms through which remote working enhanced their productivity: By reducing the social distractions of the office and by eliminating the exigencies of commuting. Regarding the social distractions of the office, Sally, a 31-year-old personal assistant, concisely summarized the consensus view: “I was more productive because less people were interrupting me and asking me bits and pieces.” Arthur, a 36-year-old change manager, elaborated further:

There is no worse place in the office to be than next to the printers. People just come and say “Hello” as they’re collecting their one pages, and then you end up getting distracted quite a lot. So I have experienced a lot of that lately ... But when I’m working from home, there’s less of that, and actually, it’s easier to manage because if someone Skypes me, I can make a decision about whether I reply to that immediately or later. Whereas, when I’m in the office and someone comes and talks to me, I’m less likely to say, “Come back later,” in the politest way possible.

The habitual interruptions that are part and parcel of office life impede concentration in addition to taking time away from work, as Amanda noted: “[While remote working] I could spend a whole hour doing some say, data analysis, without having to stop every 10 minutes or 20 minutes to answer a random question, and then come back to it. And, then it takes a few minutes to get back in the zone of what you’re doing ... I think the quality of what I was doing was better, because it was more consistent.”

Subjects also explicitly linked reduced commuting time to improved performance.

Donna, a 47-year-old senior scientist working on new product development, said that “Waking up, I mean, every day at six and spending two hours in traffic to go to work, it’s really tiring and sometimes I notice that my performance is a little bit slow ... at the end of the week, I am arriving to work really, really tired.” Similarly, Sally related a benefit of not having to commute every day: “It’s hard when you’re a working mum because you feel—you don’t want to be the person clock watching and bang on four PM you’re out of the door, but you have to do that.”

Coordination costs. Some subjects reported initial challenges regarding the logistics of Skype meetings, but often noted that these were ironed out as the trial proceeded. As Oliver described, “There’s a lot of meetings that were stalled by five or ten minutes because of getting the IT ready, set up and working as expected ... Actually now, if we re-ran the trial, I think it would be loads better. Everyone knows how to add Skype to a meeting and do the invite.”

Roger, a 37-year-old information technology manager who fielded support requests during the trial, concurred with this assessment: “Yeah, some systems, it must have been the first time they had tried to use them from home, so we had a few [issues]. But, it was remarkable how after that first cycle—so in week one, week two, we saw that a lot. Week three, week four we didn’t.”

By comparison, Amelia related that “I found I was a lot more prompt for meetings, because it was a case of—I just had to dial in at the right time rather than having to, you know, walk round, find a room, grab a cup of tea on the way. So timing-wise and meeting-wise, I thought it was actually more efficient, and I didn’t have any Skype problems, which was good.”

Jason agreed: “I don’t remember there being any time when I thought I’m blocked on my work because somebody is working remotely and I can’t talk to them.” Likewise, Donna described:

I had a long conversation with a technician that was trying to set up an experiment that I asked him to do and together, working remotely, we were able to discover a new thing about this protein and go through the literature and it was exciting. And, then here also we were working at a 50-mile distance to each other, and we get to the point that we solved the

problem, so that was really good. So, I never had any kind of issue working from home.

In terms of logistics, the most common complaint that surfaced during my interviews was the absence of dual monitors when working remotely—as compared with the office, where each workstation was equipped with two monitors. A secondary concern, which was raised by employees working in sales, involved ensuring reimbursement for long distance calls made to clients via personal phones while working from home.

Overall, however, subjects appeared to be satisfied with their experience during the experiment. As Roger related:

I mean, I have said it a few times, just how well received it was. For a company like Abcam, which is like kind of, 96% of people come to work every single day and sit at their desk, I have been really impressed with how passionately people embraced it. And, you can see that there must have been some pent-up desire in a lot of places to do this. I think one of the guys was saying, I think he went over to our Unit 200 building, which is where most of the marketing team work, and he was like, ‘There was no one there, they were all at home,’ and he saw that as a really good thing.

Likewise, Marshall, a 48-year-old executive, stated:

My own experience, I think as I mentioned to you, was I was more engaged, more productive, less fatigued and the quality of my work was better and I think my observations around my team was exactly the same, not surprisingly. So, I think what it did do was bust a number of myths for me that actually people wouldn’t work or that they needed my close supervision ... And, the other thing was this concept that you’ve got to have one-to-one [meetings] face-to-face was just a complete and utter myth ... because as I mentioned, I think it’s perfectly normal to [meet] with people in Boston or in Shanghai on Skype, so why would I assume it’s not okay to do that in the U.K.?

DISCUSSION

Scholars have spent decades diagnosing gender-based disparities in professional attainment. This effort has produced a comprehensive list of antecedents but few potential solutions. And the solutions that have been put into practice do not seem to be working. The general consequence of reduced-hours flexibility policies, for example, is to maintain women’s labor market attachment at the cost of their achievement (Mandel and Semyonov 2006, Williams et al. 2013). Broader discussions of ensuring equal pay for equal work, meanwhile, obscure a larger issue: Given the inequitable provision of childcare, equal work is, in practice, not

actually an option for many women (Kleven et al. 2018). This stark reality motivates an academic imperative: Redesigning organizations to reduce the inequality that persists within their ranks (Correll et al. 2014, Ranganathan and Pedulla 2018). The practice of discretionary remote working, while hardly a panacea in and of itself, may represent a promising step in this direction, specifically with respect to mothers working in white-collar firms.

Accordingly, in this paper I pursued three primary goals: First, to evaluate whether employee subgroups selected into remote working days at different rates when offered the opportunity to do so; second, to examine the impact of remote working with respect to the work-family interface and job performance; and third, to assess potential coordination costs, which could conceivably preclude organizations from implementing discretionary remote working as a long-term policy. To accomplish this, I conducted a randomized field experiment with a repeated cross-over design. This methodology—which is readily imitable and well-suited for randomized trials conducted outside the laboratory—was necessary because of differential selection into ad hoc remote working arrangements conditional on human and social capital, which confounds the interpretation of remote working’s impact (Weeden 2005, Davis and Kalleberg 2006, Hornung et al. 2008, Wharton et al. 2008, Golden 2008).

I can report the following results, some of which constitute contributions to the work-family and labor market inequality literatures. First, there was no significant difference in the uptake of remote working days during experimental weeks across men, women, parents, non-parents, mothers or fathers. It also bears noting that, while some comparisons approached marginal levels of statistical significance, they were not meaningfully different: All employee subgroups, on average, selected into about two days of remote working per week. This constitutes a contribution because the career penalties associated with reduced-hours flexibility policies are sustained, at least in part, by their highly gendered usage patterns

(Williams et al. 2013). Accordingly, by making remote working discretionary and distinct from reduced-hours policies, it may be possible to decouple the practice from stigmatizing associations with sex and motherhood. To be clear, it would be premature to conclude from this evidence that remote working does not engender career penalties. My results are, however, consistent with Glass and Noonan (2016), who found that remote working during regular business hours did not constrain salary growth for men or women—whereas, compared with men, mothers were disproportionately penalized for work hour reductions.

The revealed preference of two remote working days per week that the experiment elicited is also worth noting, in light of the Chinese call center results reported by Bloom and colleagues (2015). In their experiment, which evaluated the effects of working remotely for four days per week, subjects reported feeling lonely and socially isolated; as a result, a significant portion of them elected not to participate in the remote working program when it was later rolled out across the entire company. Likewise, my interview subjects were united in their view that working remotely for four or five days per week would be unappealing, largely because they would miss the social interaction afforded by the office—though they did, of course, find a moderate amount of remote working, at their own discretion, highly valuable.

Second, I found that while the effects of remote working were positive overall, mothers benefited the most. Specifically, remote working reduced family-work conflict for mothers, but not fathers or non-parents. Remote working also improved job performance for most employees, but mothers saw by far the largest gain. To contextualize these findings, I compared the mean levels of family-work conflict reported by mothers during treatment weeks to the mean levels reported by non-mothers, both male and female, during control weeks: These means were effectively equivalent. In other words, the effect of remote working was to level the playing field between mothers and other employees in terms of

family-work conflict. Mothers also reported the greatest job performance of any subgroup in the trial during remote working weeks. This hints at the possibility that remote working could “unlock” productivity in mothers that would be otherwise contained by their disproportionate caregiving duties. It is also worth noting in light of the evidence that women, particularly educated women, increasingly underestimate the effect that motherhood will have on their future labor supply (Kuziemko et al. 2018), notwithstanding a general awareness that professional success requires tradeoffs not demanded of men (Gino et al. 2015).

Of course, these effects would mean little if remote working, as a policy, could not be feasibly implemented by organizations in the long run. Impediments to feasibility include coordination costs with respect to work interdependence and informal resistance on the part of employees. The latter is not a trivial concern, given that programs designed to ameliorate inequities for a particular demographic group may engender backlash on the part of the majority (Dobbin et al. 2015). This did not appear to be the case with respect to remote working, however. For example, my interview subjects did not interpret the policy as an explicit attempt to address the professional challenges that women in general and mothers in particular face. In addition, non-parents actually reported the greatest increase in job satisfaction as a result of remote working. Despite its targeted impact, remote working appears to be a policy with universal appeal.

The evidence regarding coordination costs is slightly more nuanced. Non-parents reported a decrease in co-worker helping behavior as a result of remote working during the first half of the experiment, though parents did not. This effect receded during the second half, however, consistent with interview subjects’ assessment that it took a week to acclimate to the technological particulars of remote working. Somewhat more concerning was the fact that the performance benefits of remote working were attenuated for workers whose levels of job interdependence was above the median. This did not apply to mothers, however, who

accrued performance benefits regardless of how interdependent their role was.

Lastly, one non-finding warrants mentioning: I did not observe an increase—or decrease—in overall work hours as a result of remote working. This is important to note, as work-family scholars sometimes caution that attempts to increase employees' control over the time and place of their work could have unintended consequences vis-à-vis the amount of work that they are compelled to do (Schieman 2013). This prediction was not supported by evidence from the Abcam experiment, however.

Notwithstanding its contributions, this research is subject to several notable limitations. First and foremost, it would be reckless to extrapolate long-term effects on the basis of a four-week trial. A permanent remote working policy could, for example, result in a more traditionally gendered uptake of remote working days, while the absence of predictable in-office “control” weeks and a predetermined trial end date could produce a greater propensity for coordination failure. Clearly, further longitudinal study is required in order to determine whether the patterns that I observed would hold for a firm that made remote working continually available to all its employees. Such further study is feasible, however, in light of the fact that Abcam altered its company policy after the experiment by allowing all eligible employees to work remotely at their discretion.

Second, my results cannot generalize beyond white-collar workplaces that are similar in character to Abcam. This constitutes a meaningful scope condition with respect to remote working because the practice requires managers to place an inordinate amount of trust in their workers. One cannot presume, *ex ante*, that this trust will always be rewarded—particularly within a workforce that is less intrinsically motivated and professionally engaged than is typically the case for knowledge workers. Likewise, it bears noting that the kind of firm that participates in a field experiment is, by definition, somewhat distinct from the total population of firms—just as the Abcam employees who participated in my study may be

systematically different from those who did not. These potential confounds are a necessary feature of field experiments specifically and organizational research more generally, but it is important to remain cognizant of them when considering the generalizability of results.

My reliance on self-report measures constitutes a third limitation. While this was less of a concern for family-work conflict and job satisfaction—for which self-report is generally considered sufficient—it would have been ideal to measure job performance unobtrusively. This is, however, not a straightforward proposition for white-collar work output, which is often difficult to quantify (Goldin 2014).¹⁷ With that being said, there are professions— notably law—in which billable hours constitute the most important measure of productivity (Briscoe and Kellogg 2011, Azmat and Ferrer 2017). Law firms may therefore be an ideal setting in which to further examine the performance implications of remote working.

Across industries and professions, a dramatic revision of work practices is necessary before long-awaited progress can be made. Determining the form that these revisions should take must be an imperative for future organizational scholarship.

¹⁷ Abcam did report significant revenue growth during the period in which the trial ran, suggesting that remote working did not damage firm-level financial performance.

REFERENCES

- Acker J (1990) Hierarchies, jobs, bodies: A theory of gendered organizations. *Gend. Soc.* 4(2):139-158.
- Afflerback S, Carter SK, Anthony AK, Grauerholz L (2013) Infant feeding consumerism in the age of intensive mothering and risk society. *J. Consumer Cult.* 13(3):387-405.
- Azmat G, Ferrer R (2017) Gender gaps in performance: Evidence from Young Lawyers. *J. Political Econ.* 125(5):1306-1355.
- Bailyn L (1993 [2006]) *Breaking the Mold: Redesigning Work for Productive and Satisfying Lives* (Cornell University Press, Ithaca).
- Baldassarri D, Abascal M (2017) Field experiments across the social sciences. *Ann. Rev. Sociol.* 43:41-73.
- Bardasi E, Gornick JC (2008) Working for less? Women's part-time wage penalties across countries. *Fem. Econ.* 14(1):37-72.
- Becker PE, Moen P (1999) Scaling back: Dual-earner couples' work-family strategies. *J. Marriage Fam.* 61(4):995-1007.
- Benard S, Paik I, Correll SJ (2007) Cognitive bias and the motherhood penalty. *Hastings Law J.* 59:101-129.
- Bertrand M, Mullainathan S (2004) Are Emily and Greg more employable than Lakisha and Jamal? *Am. Econ. Rev.* 94:991-1013.
- Bertrand M, Goldin C, Katz LF (2010) Dynamics of the gender gap for young professionals in the financial and corporate sectors. *Am. Econ. J.* 2(3):228-255.
- Bianchi SM, Robinson JP, Milke M (2006) *The Changing Rhythms of American Family Life* (Russell Sage, New York).
- Blair-Loy M. (2003 [2009]) *Competing Devotions: Career and Family Among Women Executives* (Harvard University Press, Cambridge).
- Blair-Loy M. (2009) Work without end? Scheduling flexibility and work-to-family conflict among stockbrokers. *Work Occup.* 36(4):279-317.
- Blair-Loy M, Wharton AS (2002) Employees' use of work-family policies and the workplace social context. *Soc. Forces* 80(3):813-845.
- Blair-Loy M, Cech E (2014) Consequences of flexibility stigma among academic scientists and engineers. *Work Occup.* 41(1):86-110.
- Blalock AB, Blalock HM (1982) *Introduction to Social Research* (Prentice-Hall, Englewood Cliffs).
- Bloom N, Liang J, Robert J, Ying ZJ (2015) Does working from home work? Evidence from a Chinese experiment. *Q. J. Econ.* 130(1):165-218.
- Broockman DE, Kalla JL, Sekhon JS (2017) The design of field experiments with survey outcomes: A framework for selecting more efficient, robust, and ethical designs. *Political Anal.* 25(4):435-464.
- Briscoe F (2007) From iron cage to iron shield? How bureaucracy enables temporal flexibility for professional service workers. *Organ. Sci.* 18(2):297-314.
- Briscoe F, Kellogg KC (2011) The initial assignment affect: Local employer practices and positive career outcomes for work-family program users. *Am Sociol. Rev.* 76(2):291-319.
- Camman C, Fichman M, Jenkins GD, Klesh J (1983) Michigan organizational assessment questionnaire. Seashore SE, Lawler EE, Mirvis PH, Camman C, eds. *Assessing Organizational Change: A Guide to Methods, Measures, and Practices* (Wiley-Interscience, New York), 71-138.
- Cha Y (2010) Reinforcing separate spheres: The effect of spousal overwork on men's and women's employment in dual-earner households. *Am Sociol. Rev.* 75(2):303-329.
- Cha Y, Weeden KA (2014) Overwork and the slow convergence in the gender gap in wages.

- Am Sociol. Rev.* 79(3):457-484.
- Chan CK, Anteby M (2016) Task segregation as a mechanism for within-job inequality: Women and men of the transportation security administration. *Admin. Sci. Quart.* 61(2): 184-216.
- Chan RW, Brooks RC, Raboy B, Patterson CJ (1998) Division of labor among lesbian and heterosexual parents: Associations with child's adjustment. *J. Fam. Psychol.* 12(3): 402-419.
- Cohen J (1988) *Statistical Power Analysis for the Behavioral Sciences* (Academic Press, San Diego).
- Connelly BL, Tihanyi L, Crook TR, Gangloff KA (2014) Tournament theory: Thirty years of contests and competitions. *J. Manag.* 40(1):16-47.
- Correll SJ, Benard S (2006) Biased estimators? Comparing status and statistical theories of gender discrimination. *Advances in Group Processes* 23:89-116.
- Correll SJ, Benard S, Paik I (2007) Getting a job: Is there a motherhood penalty? *Am. J. Sociol.* 112(5):1297-1338.
- Correll SJ, Kelly EL, O'Connor LT, Williams JC (2014) Redesigning, redefining work. *Work Occup.* 41(1):3-17.
- Coser LA (1974) *Greedy Institutions: Patterns of Undivided Commitment* (Free Press, New York).
- Cotter D, Hermsen JM, Vanneman R (2011) The end of the gender revolution? Gender role attitudes from 1977 to 2008. *Am. J. Sociol.* 117(1):259-289.
- Cox DR (1958) *Planning of Experiments* (John Wiley & Sons, New York).
- Davis AE, Kalleberg AL (2006) Family-friendly organizations? Work and family programs in the 1990s. *Work Occup.* 33(2):191-223.
- Dekas KH, Bauer TN, Welle B, Kurkoski J, Sullivan S (2013) Organizational citizenship behavior, version 2.0: A review and qualitative investigation of OCBs for knowledge workers at Google and beyond. *Acad. Manag. Perspect.* 27(3):219-237.
- Edin K, Kefalas M (2011) *Promises I Can Keep: Why Poor Women Put Motherhood Before Marriage* (University of California Press, Oakland).
- Elliott S, Powell R, Brenton J (2015) Being a good mom: Low-income, black single mothers negotiate intensive mothering. *J. Fam. Issues* 36(3):351-370.
- England P, Bearak J, Budig MJ, Hodges MJ (2016) Do highly paid, highly skilled women experience the largest motherhood penalty? *Am Sociol. Rev.* 81(6):1161-1189.
- Epstein CF, Seron C, Oglensky B, Saute R (1999) *The Part Time Paradox: Time Norms, Professional Lives, Family and Gender* (Routledge, New York).
- Fisher RA (1935) *The Design of Experiments* (Hafner, New York).
- Gaddis SM (2015) Discrimination in the credential society: An audit study of race and college selectivity in the labor market. *Soc. Forces* 93(4):1451-1479.
- Gajendran RS, Harrison DA (2007) The good, the bad, and the unknown about telecommuting: Meta-analysis of psychological mediators and individual consequences. *J. Appl. Psychol.* 92(6):1524-1541.
- Galinsky E, Bond JT, Friedman DE (1996) The role of employers in addressing the needs of employed parents. *J. Soc. Issues* 52(3):111-136.
- Gareis KC, Barnett RC (2002) Under what conditions do long work hours affect psychological distress? *Work Occup.* 29(4):483-497.
- Gatrell CJ (2007) Secrets and lies: Breastfeeding and professional paid work. *Soc. Sci. Med.* 65(2):393-403.
- Gino F, Wilmoth CA, Brooks AW (2015) Compared to men, women view professional advancement as equally attainable, but less desirable. *Proc. Natl. Acad. Sci. U.S.A.* 112(40): 12354-12359.

- Glass J (2004) Blessing or curse? Work-family policies and mother's wage growth over time. *Work Occup.* 31(3):367-394.
- Glass JL, Finley A (2002) Coverage and effectiveness of family-responsive workplace policies. *Hum. Res. Manag. Rev.* 12(3):313-337.
- Glass JL, Noonan MC (2016) Telecommuting and earnings trajectories among American women and men 1989-2008. *Soc. Forces* 95(1):217-250.
- Glavin P, Schieman S, Reid S (2011) Boundary-spanning work demands and their consequences for guilt and psychological distress. *J. Health Soc. Behav.* 52(1):43-57.
- Golden L (2008) Limited access: Disparities in flexible work schedules and work-at-home. *J. Fam. Econ. Issues* 29:86-109.
- Goldin C (2014) A grand gender convergence: Its last chapter. *Am. Econ. Rev.* 104(4):1091-1119.
- Goldin C, Katz LF (2016) A most egalitarian profession: Pharmacy and the evolution of a family-friendly occupation. *J. Labor Econ.* 34(3):705-746.
- Greenwald AG (1976) Within-subjects designs: To use or not to use? *Psychol. Bull.* 83(2):314-320.
- Grzywacz JG, Carlson DS, Shulkin S (2008) Schedule flexibility and stress: Linking formal flexible arrangements and perceived flexibility to employee health. *Comm. Work Fam.* 11(2):199-214.
- Hays S (1996) *The Cultural Contradictions of Motherhood* (Yale University Press, New Haven).
- Hewlett SA (2002) *Creating a Life: Professional Women and the Quest for Children* (Talk Maramax Books, New York).
- Hills M, Armitage P (1979) The two-period cross-over clinical trial. *Brit. J. Clinical Pharm.* 8:7-20.
- Hollis S, Campbell F (1999) What is meant by intention to treat analysis? Survey of published randomised controlled trials. *BMJ* 319(7211):670-674.
- Hornung S, Rousseau DM, Glaser J (2008) Creating flexible work arrangements through idiosyncratic deals. *J. Appl. Psychol.* 93(3):655-664.
- Jackson M, Cox DR (2013) The principles of experimental design and their application in sociology. *Ann. Rev. Sociol.* 39:27-49.
- Jacobs JA, Gerson K (2004) *The Time Divide: Work, Family, and Gender Inequality* (Harvard University Press, Cambridge).
- Jones B, Kenward MG (2015) *Design and Analysis of Cross-over Trials* (Taylor & Francis Group, London).
- Kahn JR, Garcia-Mangano J, Bianchi SM (2014) The motherhood penalty at midlife: Long-term effects of children on women's careers. *J. Marriage Fam.* 76(1):56-72.
- Kalleberg AL, Reskin B (1995) Gender differences in promotion in the United States and Norway. *Res. Soc. Stratif. Mobil.* 14:237-264.
- Kelly EL, Ammons SK, Chermack K, Moen P (2010) Gendered challenge, gendered response: Confronting the ideal worker norm in a white-collar organization. *Gend. Soc.* 24(3):281-303.
- Kelly EL, Moen P, Oakes JM, Fan W, Okechukwu C, Davis KD, Hammer LB, Kossek EE, King RB, Hanson GC, Mierzwa F, Casper LM (2014) Changing work and work-family conflict: Evidence from the work, family, and health network. *Am Sociol. Rev.* 79(3):485-516.
- Killewald A, Garcia-Mangano J (2016) Tethered lives: A couple-based perspective on the consequences of parenthood for time use, occupation, and wages. *Soc. Sci. Res.* 60:266-282.
- Klevin HJ, Landais C, Sogaard JE (2018) Children and gender inequality: Evidence from

- Denmark. Working paper.
- Kossek EE, Lautsch BA, Eaton SC (2006) Telecommuting, control, and boundary management: Correlates of policy use and practice, job control, and work-family effectiveness. *J. Vocat. Behav.* 68:347-367.
- Kossek EE, Lautsch BA (2018) Work-life flexibility for whom? Occupational status and work-life inequality in upper, middle, and lower level jobs. *Acad. Manag. Annals* 12(1):5-36.
- Kuziemko I, Pan J, Shen J, Washington E (2018) The mommy effect: Do women anticipate the employment effects of motherhood? NBER Working Paper No. 24740.
- Lamb ME, Tamis-Lemonda CS (2004) The Role of the Father: An Introduction. Lamb, ME ed. *The Role of the Father* (John Wiley & Sons, Hoboken), 1-31.
- Lareau A (2000) *Home Advantage: Social Class and Parental Intervention in Elementary Education* (Rowman & Littlefield, Lantham).
- Lareau A (2002) Invisible inequality: Social class and childrearing in black families and white families. *Am Sociol. Rev.* 67(5):747-776.
- Levitt SD, List JA (2009) Field experiments in economics: The past, the present, and the future. *Eur. Econ. Rev.* 53(1):1-18.
- Lyons K, Lockie S, Lawrence G (2001) Consuming 'green': The symbolic construction of organic foods. *Rural Sociol.* 11(3):197-210.
- Mandel H, Semyonov M (2006) A welfare state paradox: State interventions and women's employment opportunities in 22 countries. *Am. J. Sociol.* 111(6):1910-1949.
- Martens CR, Denman BA, Mazzo MR, Armstrong ML, Reisdorph N, McQueen MB, Chonchol M, Seals DR (2018) Chronic nicotinamide riboside supplementation is well-tolerated and elevates NAD+ in healthy middle-aged and older adults. *Nat. Commun.* 9(1286):1-11.
- Maume DJ (2006) Gender differences in restricting work efforts because of family responsibilities. *J. Marriage Fam.* 68(4):859-869.
- Maume DJ (2008) Gender differences in providing urgent childcare among dual-earner parents. *Soc. Forces* 87(1):273-297.
- Maume DJ, Sebastian RA, Bardo AR (2009) Gender differences in sleep disruption among retail workers. *Am Sociol. Rev.* 74(6):989-1007.
- Maume DJ, Sebastian RA, Bardo AR (2010) Gender, work-family responsibilities, and sleep. *Gend. Soc.* 24(6):746-768.
- McPherson D (1993) *Gay Parenting Couples: Parenting Arrangements, Arrangement Satisfaction, and Relationship Satisfaction*. Unpublished doctoral dissertation.
- Meyers M, Gornick J (2005) Policies for reconciling parenthood and employment: Drawing lessons from Europe. *Challenge* 48:39-61.
- Miller CC (2017) How to close the gender gap: Let employees control their schedules. *The New York Times*, February 7.
- Mishel L (2013) Vast majority of wage earners are working harder, and for not much more: Trends in U.S. work hours and wages 1979-2007. Issue Brief No. 348. Washington, DC: Economic Policy Institute.
- Moen P, Kelly EL, Fan W, Lee S, Almeida D, Kossek EE, Buxton OM (2016) Does a flexibility/support organizational initiative improve high-tech employees' well-being? Evidence from the work, family, and health network. *Am Sociol. Rev.* 81(1):134-164.
- Moisio R, Arnould EJ, Price LL (2004) Between mothers and markets: Constructing family identity through homemade food. *J. Consumer Cult.* 4(3):361-384.
- Morgan SL, Winship C (2007 [2014]) *Counterfactuals and Causal Inference* (Cambridge University Press, New York).
- Nelson MK (2010) *Parenting Out of Control: Anxious Parents in Uncertain Times* (NYU

- Press, New York).
- Netermeyer RG, Boles JS, McMurrian R (1996) Development and validation of work-family conflict and family-work conflict scales. *J. Appl. Psychol.* 81(4):400-410.
- Noonan MC (2001) The impact of domestic work on men's and women's wages. *J. Marriage Fam.* 63(4):1134-1145.
- Nyhan B, Reifler J (2015) The effect of fact-checking on elites: A field experiment on U.S. state legislators. *Am. J. Political Sci.* 59(3):628-640.
- Offer S, Schneider B (2011) Revisiting the gender gap in time-use patterns: Multitasking and well-being among mothers and fathers in dual-earner families. *Am Sociol. Rev.* 76(6):809-833.
- Orne MT (1973) Communication by the total experimental situation: Why it is important, how it is evaluated, and its significance for the ecological validity of findings. Pliner P, Krames L, Alloway TM, eds. *Communication and Affect* (Academic Press, New York), 157-191).
- Pager D (2003) The mark of a criminal record. *Am. J. Sociol.* 108(5):937-975.
- Pedulla DS, Thebaud S (2015) Can we finish the revolution? Gender, work-family ideals, and institutional constraint. *Am Sociol. Rev.* 80(1):116-139.
- Petersen T, Penner AM, Høgsnes G (2014) From motherhood penalties to husband premia: The new challenge for gender equality and family policy, lessons from Norway. *Am. J. Sociol.* 119(5):1434-1472.
- Prentice DA, Miller DT (1992) When small effects are impressive. *Psychol. Bull.* 112(1):160-164.
- Quadlin N (2018) The mark of a woman's record: Gender and academic performance in hiring. *Am Sociol. Rev.* forthcoming.
- Ranganathan A, Pedulla D (2018) Organizations as network equalizers? Employer-provided childcare and the labor supply of working mothers. Working paper.
- Ridgeway CL, Correll SJ (2004) Motherhood as a status characteristic. *J. Soc. Issues* 60(4):683-700.
- Rippeyoung PLF, Noonan MC (2012) Is breastfeeding truly cost-free? Income consequences of breastfeeding for women. *Am Sociol. Rev.* 77(2):244-267.
- Rivera LA, Tilcsik A (2017) Class advantage, commitment penalty: The gendered effect of social class signals in an elite labor market. *Am Sociol. Rev.* 81(6):1097-1131.
- Roethlisberger FJ, Dickson WJ (1939 [2003]) *Management and the Worker* (Harvard University Press, Cambridge).
- Rousseau DM (2006) Is there such a thing as 'evidence-based management'? *Acad. Manag. Rev.* 31(2):256-269.
- Rousseau DM, Gunia B (2016) Evidence-based practice: The psychology of EBP Implementation. *Ann. Rev. Psychol.* 67:667-692.
- Ruch FL, Zimbardo PG (1971) *Psychology and Life, 8th Edition* (Scott Foresman, Glenview).
- Sanchez L, Thomson E (1997) Becoming mothers and fathers: Parenthood, gender, and the division of labor. *Gend. Soc.* 11(6):747-772.
- Sayer LC, Bianchi SM, Robinson JP (2004) Are parents investing less in children? Trends in mothers' and fathers' time with children. *Am. J. Sociol.* 110(1):1-43.
- Schieman S (2013) Job-related resources and the pressures of working life. *Soc. Sci. Res.* 42:271-282.
- Sears W, Sears M (1992) *The Baby Book: Everything You Need to Know About Your Baby—From Birth to Age Two* (Little, Brown & Co., New York).
- Sewell WH (1992) A theory of structure: Duality, agency, and transformations. *Am. J. Sociol.* 98(1):1-29.
- Sharone O (2004) Engineering Overwork: Bell Curve Management at a High-Tech Firm.

- Epstein CF, Kalleberg AL, eds. *Fighting for Time: Shifting Boundaries of Work and Social Life* (Russell Sage Foundation, New York), 191-218.
- Simmons JP, Nelson LD, Simonsohn U (2011) False-positive psychology: Undisclosed flexibility in data collection and analysis allows presenting anything as significant. *Psychol. Sci.* 22(11):1359-1366.
- Simon RW (1995) Gender, multiple roles, role meaning, and mental health. *J. Health Soc. Behav.* 36(2):182-194.
- Stone P (2008) *Opting Out? Why Women Really Quit Careers and Head Home* (University of California Press, Oakland).
- Stone P, Hernandez LA (2013) The all-or-nothing workplace: Flexibility stigma and 'opting out' among professional-managerial women. *J. Soc. Issues* 69(2):235-256.
- Sullivan O (1997) Time waits for no (wo)man: An investigation of the gendered experience of domestic time. *Sociology* 31(2):221-239.
- Thebaud S, Pedulla DS (2016) Masculinity and the stalled revolution: How gender ideologies and norms shape young men's responses to work-family policies. *Gend. Soc.* 30(4):590-617.
- Tilcsik A (2011) Pride and prejudice: Employment discrimination against openly gay men in the United States. *Am. J. Sociol.* 117(2):586-626.
- Turco CJ (2010) Cultural foundations of tokenism. *Am Sociol. Rev.* 75(6):894-913.
- Van der Veegt GS, Evans BJM, Van De Vliert E (2001) Patterns of interdependence in work teams: A two-level investigation of the relations with job and team satisfaction. *Pers. Psychol.* 54:51-69.
- Voydanoff P (2004) The effects of work demands and resources on work-to-family conflict and facilitation. *J. Marriage Fam.* 66(2):398-412.
- Voydanoff P (2007) *Work, Family, and Community: Exploring Interconnections* (Lawrence Erlbaum Associates, Mahwah).
- Vincent C, Ball SJ (2007) 'Making up' the middle class child: Families, activities, and class dispositions. *Sociology* 41(6):1061-1077.
- Webber G, Williams C (2008) Mothers in 'good' and 'bad' part-time jobs: Different problems, same results. *Gend. Soc.* 22(6):752-777.
- Weeden KA (2005) Is there a flexiglass ceiling? Flexible work arrangements and wages in the United States. *Soc. Sci. Res.* 34:454-482.
- Weisshaar K (2018) From opt out to blocked out: The challenges of labor market re-entry after family-related employment lapses. *Am Sociol. Rev.* 83(1):34-60.
- Welbourne TM, Johnson DE, Erez A (1998) The Role-based performance scale: Validity analysis of a theory-based measure. *Acad. Manag. J.* 41(5):540-555.
- Wharton AS, Chivers S, Blair-Loy M (2008) Use of formal and informal work family policies on the digital assembly line. *Work Occup.* 35(3):327-350.
- Wilde ET, Batchelder L, and Ellwood DT (2010) The mommy track divides: The impact of childbearing on wages of women of differing skill levels. NBER Working Paper No. 16582.
- Williams JC (2001) *Unbending Gender: Why Family and Work Conflict and What to do About it* (Oxford University Press, New York).
- Williams JC (2010) *Reshaping the Work-Family Divide: Why Men and Class Matter* (Harvard University Press, Cambridge).
- Williams JC, Blair-Loy M, Berdahl JL (2013) Cultural schemas, social class, and the flexibility stigma. *J. Soc. Issues* 69(2):209-234.
- Zizzo DJ (2010) Experiment demand effects in economic experiments. *Exp. Econ.* 13(1):75-98.

Table 1 – Descriptive Statistics

| Variable | Min. | Max. | Mean | S.D. |
|---------------------------------------|-------------|-------------|-------------|-------------|
| Age (in years) | 19 | 65 | 34.10 | 8.16 |
| Female | 0 | 1 | 0.52 | 0.50 |
| Mother | 0 | 1 | 0.18 | 0.38 |
| Father | 0 | 1 | 0.19 | 0.39 |
| At least one child in household | 0 | 1 | 0.37 | 0.48 |
| Annual salary (in pounds sterling) | 15,000 | 170,000 | 44,730.01 | 21,441.48 |
| Manager | 0 | 1 | 0.39 | 0.49 |
| Living with spouse | 0 | 1 | 0.44 | 0.50 |
| Living with partner | 0 | 1 | 0.29 | 0.46 |
| Living with housemates | 0 | 1 | 0.14 | 0.35 |
| Living alone | 0 | 1 | 0.12 | 0.33 |
| Weekly work hours at baseline | 0 | 60 | 38 | 6.28 |
| Remote working days at baseline | 0 | 5 | 0.33 | 0.63 |
| Family-to-work conflict at baseline | 1 | 4.8 | 1.78 | 0.84 |
| Job performance at baseline | 2 | 5 | 4.05 | 0.57 |
| Job satisfaction at baseline | 1 | 5 | 3.97 | 0.80 |
| Helping behavior at baseline | 2.75 | 5 | 4.20 | 0.50 |
| Job interdependence at baseline | 1.4 | 4.6 | 3.12 | 0.56 |
| Flex hours at baseline | 0 | 20 | 2.17 | 3.33 |
| <i>Pooled Control Week Measures</i> | | | | |
| Weekly work hours | 0 | 75 | 36.18 | 8.90 |
| Flex hours | 0 | 19 | 1.37 | 2.80 |
| Remote working days | 0 | 5 | 0.49 | 0.93 |
| Family-to-work conflict | 1 | 5 | 1.64 | 0.86 |
| Job performance | 1 | 5 | 3.98 | 0.61 |
| Job satisfaction | 1 | 5 | 3.80 | 0.84 |
| Helping behavior | 1 | 5 | 3.93 | 0.61 |
| <i>Pooled Treatment Week Measures</i> | | | | |
| Weekly work hours | 0 | 60 | 36.47 | 9.27 |
| Flex hours | 0 | 42 | 3.04 | 6.04 |
| Remote working days | 0 | 5 | 2.14 | 1.54 |
| Family-to-work conflict | 1 | 5 | 1.60 | 0.84 |
| Job performance | 2 | 5 | 4.10 | 0.59 |
| Job satisfaction | 1 | 5 | 3.91 | 0.84 |
| Helping behavior | 1 | 5 | 3.81 | 0.72 |

N = 187 participants, 748 participant-weeks

Note: Subjects can work zero hours in a week due to vacation time.

Table 2 – Uptake of Remote Working Days per Week at Baseline Across Subsamples

| Sample restriction | Min. | Max. | Mean | S.D. | T-test |
|---------------------------|-------------|-------------|-------------|-------------|---------------|
| Full sample | 0 | 5 | 0.33 | 0.63 | N/A |
| Men | 0 | 3 | 0.32 | 0.59 | $p = 0.77$ |
| Women | 0 | 5 | 0.35 | 0.68 | $p = 0.77$ |
| Parents | 0 | 5 | 0.37 | 0.73 | $p = 0.56$ |
| Non-parents | 0 | 3 | 0.31 | 0.57 | $p = 0.56$ |
| Mothers | 0 | 5 | 0.44 | 0.94 | $p = 0.30$ |
| Fathers | 0 | 2 | 0.31 | 0.47 | $p = 0.76$ |

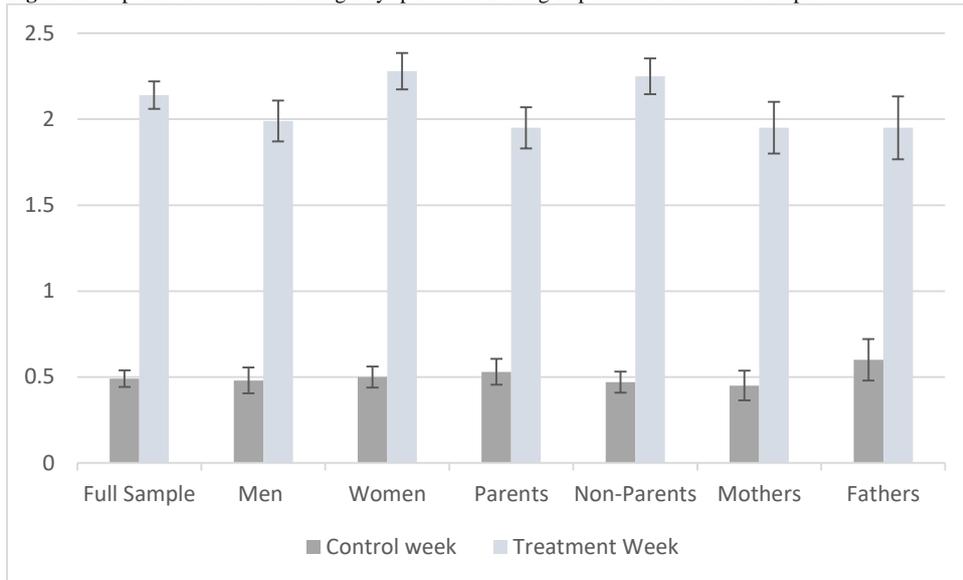
Note: T-tests indicate comparisons between the focal subgroup and the rest of the sample.

Table 3 – Uptake of Remote Working Days per Week During Treatment Weeks Across Subsamples

| Sample restriction | Min. | Max. | Mean | S.D. | T-test |
|---------------------------|-------------|-------------|-------------|-------------|---------------|
| Full sample | 0 | 5 | 2.14 | 1.54 | N/A |
| Men | 0 | 5 | 1.99 | 1.58 | $p = 0.06$ |
| Women | 0 | 5 | 2.28 | 1.48 | $p = 0.06$ |
| Parents | 0 | 5 | 1.95 | 1.40 | $p = 0.07$ |
| Non-parents | 0 | 5 | 2.25 | 1.60 | $p = 0.07$ |
| Mothers | 0 | 5 | 1.95 | 1.22 | $p = 0.28$ |
| Fathers | 0 | 5 | 1.95 | 1.56 | $p = 0.24$ |

Note: T-tests indicate comparisons between the focal subgroup and the rest of the sample.

Figure 1 – Uptake of Remote Working Days per Week During Experiment Across Subsamples



Note: Figure depicts a between-subjects comparison of means.

Table 4 – The Effect of Remote Working on Family-to-work Conflict by Subgroup, Within Subjects

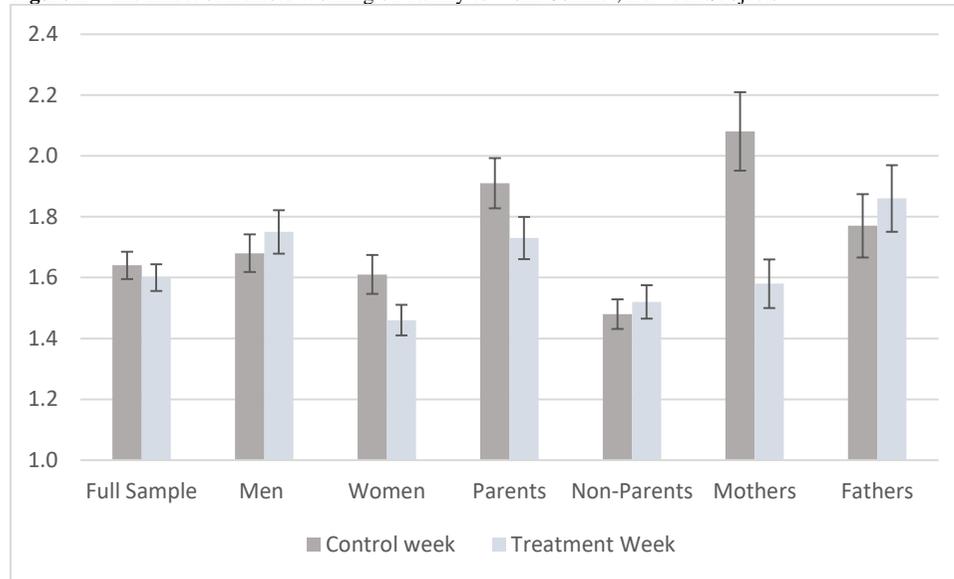
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) |
|--------------------------|-------------------------|-------------------------|-------------------------|-------------------------|-------------------------|-------------------------|-------------------------|
| | Family-work Conflict |
| Sample restriction | Full sample | Men | Women | Parents | Non-parents | Mothers | Fathers |
| Remote working week | -0.0416 (0.0487) | 0.0773 (0.0738) | -0.147* (0.0638) | -0.185+ (0.0945) | 0.0424 (0.0538) | -0.492** (0.133) | 0.0986 (0.131) |
| Individual fixed effects | Yes |
| Week fixed effects | Yes |
| Cohen's <i>d</i> | 0.05 | 0.09 | 0.19 | 0.21 | 0.05 | 0.57 | 0.11 |
| Observations | 748 | 356 | 392 | 276 | 472 | 132 | 144 |
| <i>R</i> ² | 0.002 | 0.010 | 0.029 | 0.021 | 0.003 | 0.130 | 0.010 |

Standard errors in parentheses

Cohen (1988) describes an effect size of 0.20 as small, an effect size of 0.50 as medium, and an effect size of 0.80 as large.

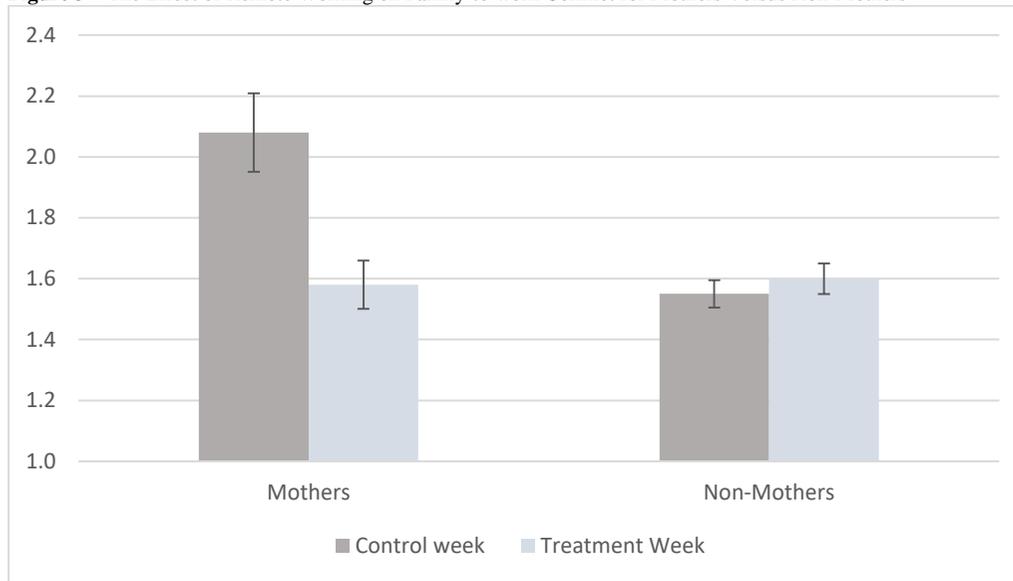
+ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$

Figure 2 – The Effect of Remote Working on Family-to-work Conflict, Between Subjects



Note: Figure depicts a pooled comparison of means.

Figure 3 – The Effect of Remote Working on Family-to-work Conflict for Mothers Versus Non-Mothers



Note: Figure depicts a pooled comparison of means.

Table 5 – The Effect of Remote Working on Job Performance, Within Subjects

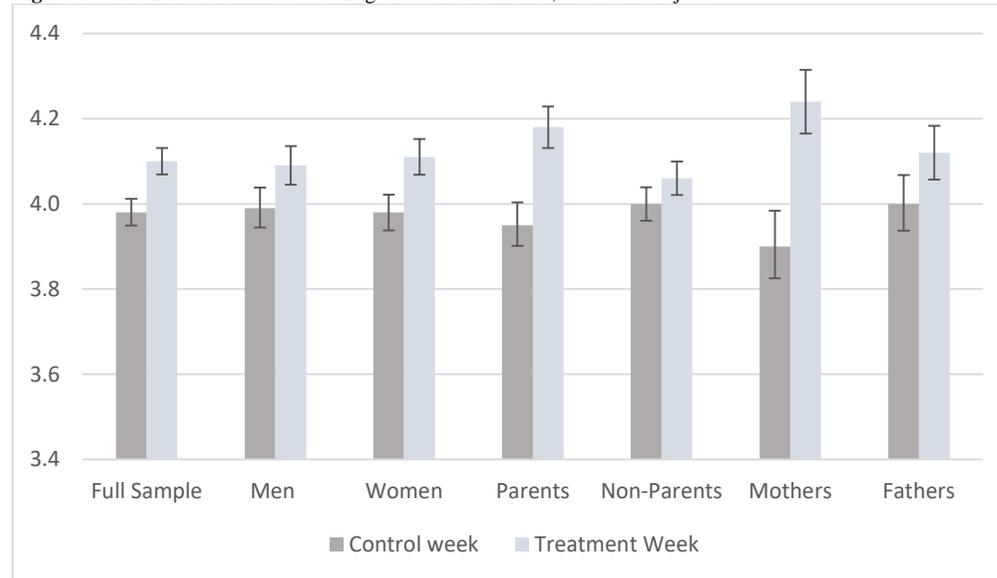
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) |
|--------------------------|---------------------|--------------------|---------------------|---------------------|--------------------|---------------------|-------------------|
| | Performance | Performance | Performance | Performance | Performance | Performance | Performance |
| Sample restriction | Full sample | Men | Women | Parents | Non-parents | Mothers | Fathers |
| Remote working week | 0.118** (0.0344) | 0.106* (0.0525) | 0.126** (0.0452) | 0.225** (0.0561) | 0.0551 (0.0431) | 0.336** (0.0826) | 0.118 (0.0765) |
| Individual fixed effects | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Week fixed effects | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Cohen's <i>d</i> | 0.20 | 0.17 | 0.22 | 0.38 | 0.09 | 0.53 | 0.22 |
| Observations | 748 | 356 | 392 | 276 | 472 | 132 | 144 |
| <i>R</i> ² | 0.023 | 0.024 | 0.039 | 0.093 | 0.018 | 0.163 | 0.056 |

Standard errors in parentheses

Cohen (1988) describes an effect size of 0.20 as small, an effect size of 0.50 as medium, and an effect size of 0.80 as large.

+ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$

Figure 4 – The Effect of Remote Working on Job Performance, Between Subjects



Note: Figure depicts a pooled comparison of means.

Table 6 – The Effect of Remote Working on Job Satisfaction by Subgroup, Within Subjects

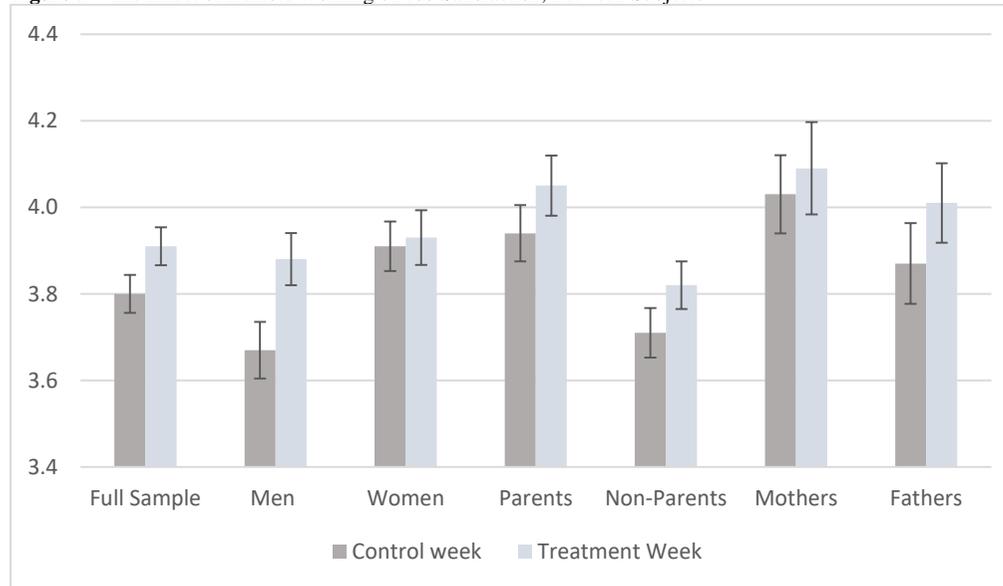
| | (1) Job Satisfaction | (2) Job Satisfaction | (3) Job Satisfaction | (4) Job Satisfaction | (5) Job Satisfaction | (6) Job Satisfaction | (7) Job Satisfaction |
|--------------------------|----------------------------|----------------------------|----------------------------|----------------------------|----------------------------|----------------------------|----------------------------|
| Sample restriction | Full sample | Men | Women | Parents | Non-parents | Mothers | Fathers |
| Remote working week | 0.107** (0.0364) | 0.203** (0.0482) | 0.0172 (0.0530) | 0.101+ (0.0546) | 0.110* (0.0480) | 0.0398 (0.0834) | 0.133+ (0.0693) |
| Individual fixed effects | Yes |
| Week fixed effects | Yes |
| Cohen's <i>d</i> | 0.13 | 0.24 | 0.02 | 0.13 | 0.13 | 0.08 | 0.18 |
| Observations | 748 | 356 | 392 | 276 | 472 | 132 | 144 |
| <i>R</i> ² | 0.029 | 0.095 | 0.029 | 0.024 | 0.042 | 0.089 | 0.062 |

Standard errors in parentheses

Cohen (1988) describes an effect size of 0.20 as small, an effect size of 0.50 as medium, and an effect size of 0.80 as large.

+ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$

Figure 5 – The Effect of Remote Working on Job Satisfaction, Between Subjects



Note: Figure depicts a pooled comparison of means.

Table 7 – The Effect of Remote Working on Co-Worker Helping Behavior by Subgroup, Within Subjects

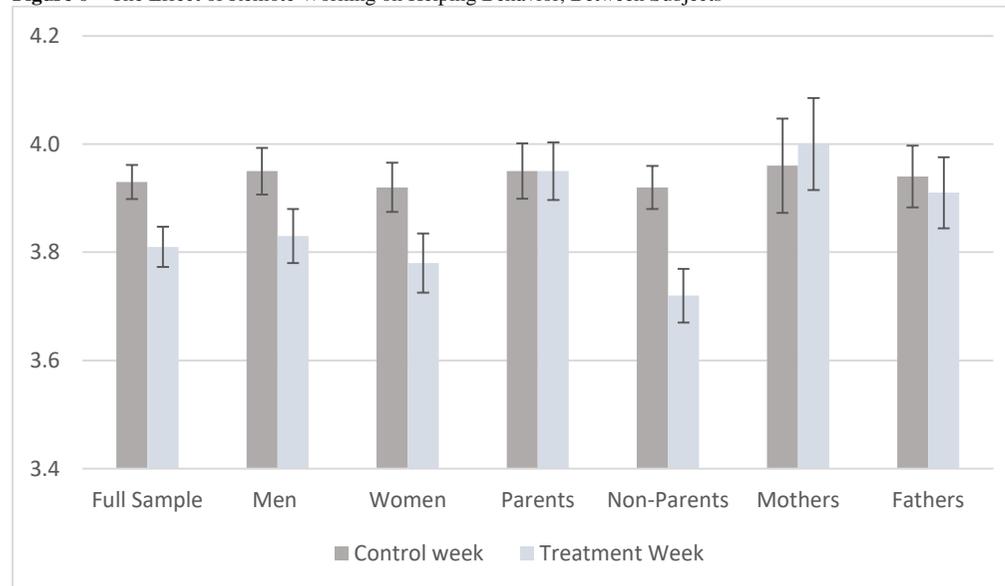
| | (1) Helping | (2) Helping | (3) Helping | (4) Helping | (5) Helping | (6) Helping | (7) Helping |
|--------------------------|----------------------|---------------------|---------------------|----------------------|----------------------|--------------------|---------------------|
| Sample restriction | Full sample | Men | Women | Parents | Non-parents | Mothers | Fathers |
| Remote working week | -0.127** (0.0414) | -0.118* (0.0578) | -0.138* (0.0590) | -0.00168 (0.0559) | -0.201** (0.0567) | 0.0250 (0.0884) | -0.0298 (0.0695) |
| Individual fixed effects | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Week fixed effects | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Cohen's <i>d</i> | 0.19 | 0.19 | 0.19 | 0.00 | 0.29 | 0.05 | 0.06 |
| Observations | 748 | 356 | 392 | 276 | 472 | 132 | 144 |
| <i>R</i> ² | 0.036 | 0.024 | 0.059 | 0.037 | 0.050 | 0.088 | 0.040 |

Standard errors in parentheses.

Cohen (1988) describes an effect size of 0.20 as small, an effect size of 0.50 as medium, and an effect size of 0.80 as large.

+ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$

Figure 6 – The Effect of Remote Working on Helping Behavior, Between Subjects



Note: Figure depicts a pooled comparison of means.

Appendix 1: Subject Recruitment Email

Abcam has been approached by London Business School to collaborate on a research project. You are eligible to participate and I strongly encourage you to do so. Your involvement will contribute to the generation of important research and, in exchange for your time, you will receive a £50 John Lewis voucher. Further details are included below.

The purpose of the project is to evaluate the effects of remote working, defined as work that is completed from outside of the office - for example at your home, or in a coffee shop. The project will run for four consecutive weeks beginning on Monday, September 25. During two of these weeks we will ask you to relax the assumption that work must necessarily be done in the office and, instead, work remotely for as much as is sensible given your professional responsibilities. The other two weeks will be business as usual: You will work in the office as much as you would normally do. You will know in advance which weeks will emphasise remote working and which weeks will consist of business as usual.

The impact of remote working will be evaluated via a short weekly survey that you will complete electronically on your phone, tablet, or computer. Each survey is very brief, requiring approximately four minutes to complete. However, there is a limited window to complete each survey: Each week the link will be e-mailed to you on Friday at noon and will only stay open through 23:59 on Sunday. **You must complete every survey in order to receive the £50 John Lewis voucher.**

Please click on the link at the bottom of this e-mail if you wish to register for the project. The link will take you to a website where you will indicate your consent to participate. You will then complete a baseline survey which is required for enrolment in the project. Please note that the baseline survey is slightly longer than the weekly surveys you will complete throughout the project period. Thank you for your consideration!

Appendix 2 – Scale Items

| Scale | Source | Items |
|-------------------------|----------------------------|--|
| Family-to-work conflict | Netemeyer et al. (1996) | The demands of my family or personal relationships interfered with work-related activities. I had to put off doing things at work because of demands on my time from home. Things I wanted to do at work didn't get done because of the demands of my family or personal life. My home life interfered with my responsibilities at work, such as getting to work on time, accomplishing daily tasks, and working overtime. Family-related strain interfered with my ability to perform job-related duties. |
| Job performance | Welbourne et al. (1998) | How would you evaluate your job performance with respect to the quantity of your work output? How would you evaluate your job performance with respect to the quality of your work output? How would you evaluate your job performance with respect to the accuracy of your work output? |
| Job satisfaction | Camman et al. (1983) | In general, I like working at my job. In general, I am satisfied with my job. I am generally satisfied with the kind of work I do in my job. |
| Co-worker helping | Dekas et al. (2013) | I helped others who had heavy workloads. I willingly helped others solve work-related problems. I was always ready to lend a helping hand to those around me. I communicated with others before initiating actions that might affect them. |
| Job interdependence | Van der Vegt et al. (2001) | I have to obtain information and advice from my colleagues in order to complete my work. I depend on my colleagues for the completion of my work. I have a one-person job; I rarely have to check or work with others. (Reverse-coded) I have to work closely with my colleagues to do my work properly. In order to complete their work, my colleagues have to obtain information and advice from me. |

All items except for job performance were measured on the following scale: 1 = strongly disagree, 2 = disagree, 3 = neither agree nor disagree, 4 = agree, 5 = strongly agree.
Job performance was measured on the following scale: 1 = needs much improvement, 2 = needs some improvement, 3 = satisfactory, 4 = good, 5 = excellent.

Appendix 3 – The Effect of Weekly Remote Working Days on Family-to-work Conflict by Subgroup, Within Subjects

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) |
|--------------------------|-------------------------|-------------------------|-------------------------|-------------------------|-------------------------|-------------------------|-------------------------|
| | Family-work Conflict |
| Sample restriction | Full sample | Men | Women | Parents | Non-parents | Mothers | Fathers |
| Remote working days | -0.00412 (0.0196) | 0.0260 (0.0298) | -0.0293 (0.0256) | -0.103* (0.0403) | 0.0429* (0.0207) | -0.214** (0.0622) | -0.0261 (0.0529) |
| Individual fixed effects | Yes |
| Week fixed effects | Yes |
| Observations | 748 | 356 | 392 | 276 | 472 | 132 | 144 |
| R ² | 0.001 | 0.008 | 0.016 | 0.034 | 0.014 | 0.115 | 0.007 |

Standard errors in parentheses.

+ p < 0.10, * p < 0.05, ** p < 0.01

Appendix 4 – The Effect of Weekly Remote Working Days on Job Performance by Subgroup, Within Subjects

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) |
|--------------------------|--------------------|---------------------|---------------------|---------------------|----------------------|---------------------|---------------------|
| | Performance | Performance | Performance | Performance | Performance | Performance | Performance |
| Sample restriction | Full sample | Men | Women | Parents | Non-parents | Mothers | Fathers |
| Remote working days | 0.0144 (0.0140) | -0.0106 (0.0214) | 0.0355+ (0.0182) | 0.0481+ (0.0248) | 0.000198 (0.0167) | 0.114** (0.0399) | 0.00141 (0.0312) |
| Individual fixed effects | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Week fixed effects | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 748 | 356 | 392 | 276 | 472 | 132 | 144 |
| R ² | 0.005 | 0.010 | 0.026 | 0.039 | 0.013 | 0.094 | 0.034 |

Standard errors in parentheses.

+ p < 0.10, * p < 0.05, ** p < 0.01

Appendix 5 – The Effect of Weekly Remote Working Days on Job Satisfaction by Subgroup, Within Subjects

| | (1) Job Satisfaction | (2) Job Satisfaction | (3) Job Satisfaction | (4) Job Satisfaction | (5) Job Satisfaction | (6) Job Satisfaction | (7) Job Satisfaction |
|--------------------------|----------------------------|----------------------------|----------------------------|----------------------------|----------------------------|----------------------------|----------------------------|
| Sample restriction | Full sample | Men | Women | Parents | Non-parents | Mothers | Fathers |
| Remote working days | 0.0548** (0.0145) | 0.0748** (0.0196) | 0.0352+ (0.0210) | 0.0774** (0.0230) | 0.0458* (0.0186) | 0.0674+ (0.0381) | 0.0804** (0.0273) |
| Individual fixed effects | Yes |
| Week fixed effects | Yes |
| Observations | 748 | 356 | 392 | 276 | 472 | 132 | 144 |
| R ² | 0.039 | 0.085 | 0.038 | 0.060 | 0.044 | 0.116 | 0.103 |

Standard errors in parentheses

+ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$

Appendix 6 – The Effect of Weekly Remote Working Days on Job Co-Worker Helping by Subgroup, Within Subjects

| | (1) Helping | (2) Helping | (3) Helping | (4) Helping | (5) Helping | (6) Helping | (7) Helping |
|--------------------------|----------------------|---------------------|---------------------|---------------------|----------------------|--------------------|----------------------|
| Sample restriction | Full sample | Men | Women | Parents | Non-parents | Mothers | Fathers |
| Remote working days | -0.0330* (0.0167) | -0.0259 (0.0235) | -0.0388 (0.0237) | 0.00924 (0.0240) | -0.0533* (0.0222) | 0.0337 (0.0409) | -0.00991 (0.0281) |
| Individual fixed effects | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Week fixed effects | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 748 | 356 | 392 | 276 | 472 | 132 | 144 |
| R ² | 0.026 | 0.013 | 0.051 | 0.038 | 0.032 | 0.094 | 0.040 |

Standard errors in parentheses

+ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$