Does Remote Work Reinforce Gender Gaps in (Un)Paid Labor?

Jean-Victor Alipour (LMU Munich, ifo Institute)

Discussion Paper No. 542

August 04, 2025

Does Remote Work Reinforce Gender Gaps in (Un)Paid Labor?

Jean-Victor Alipour[a,*]

[a] alipour@ifo.de. *University of Munich (LMU), ifo Institute for Economic Research, and CESifo*

August 4, 2025 (Click here for latest version)

Abstract

I study how the rise in working from home (WFH) affects the gender division of paid and unpaid labor (caregiving, domestic tasks). Identification uses differences in individuals' exposure to the Covid-induced WFH shock, measured by the WFH feasibility of their job in 2019. Using panel data from the German SOEP, I estimate 2SLS models that instrument realized WFH in 2022 with WFH feasibility. Results show that WFH reduces paid hours and increases domestic work and leisure (including sleep) among women. Men's time use remains largely unchanged, partly because WFH induces moves toward larger, more distant homes, offsetting commuting time savings. Within-couple analyses confirm that the Big Shift to WFH intensifies gender gaps in paid and unpaid work, particularly caregiving. I find that gender norms, bargaining power, and childcare demands interact with WFH in ways that reinforce the unequal division of labor.

Keywords: Work from home; time use; gender gaps; unpaid work; division of labor; gender norms; bargaining power

JEL Codes: J16, J22, J13

*I gratefully acknowledge funding by the DFG (Deutsche Forschungsgemeinschaft) through CRC TRR 190 (project number 280092119).

1 Introduction

Since the coordinated shift in 2020, working from home (WFH) has become a defining feature of the modern labor market (Aksoy *et al.*, 2022, 2025). Supported by network externalities, investment, and innovation in WFH technologies, employers have shed productivity concerns and extended new flexibility to a broad segment of the workforce (Barrero *et al.*, 2023, 2021; Bloom *et al.*, 2024).[1]

This dual shift, in work arrangements and perceptions, has sparked competing views on WFH's implications for gender inequality in paid and unpaid labor (caregiving, domestic tasks). Convergence optimists argue that WFH enables greater labor market participation for women and lowers barriers to men's household involvement–also thanks to time savings and the fading stigma once tied to WFH. Skeptics counter that WFH reinforces traditional gender roles by providing the pretext to burden women with more unpaid work. Men would instead capitalize on WFH to boost their productivity and careers.

This paper estimates the causal impact of WFH on the gender division of (un)paid labor and leisure in the post-Covid economy. I use representative panel data from the German Socioeconomic Panel (SOEP v.39), which includes yearly time-use information at the individual and couple level until 2022. Identification uses individual-level differences in the exposure to the WFH shock, measured by 2019 *WFH feasibility*, i.e., whether a job *could* be done from home. I estimate intention-to-treat (ITT) effects of WFH using a difference-in-differences (DiD) design that compares individuals by WFH feasibility over time. To address non-compliance with initial WFH exposure, I instrument realized WFH in 2022 with pre-Covid WFH feasibility, separately by gender. At the couple level, the IV model links changes in individual time use each partner's WFH status in 2022, instrumented by each partner's 2019 WFH feasibility.

The key takeaway is that the shift to WFH reinforces gender disparities in both paid and unpaid work. WFH reduces women's working hours and increases their time spent on leisure and domestic labor. In contrast, men's time allocation is barely affected by WFH, which can be reconciled with moves to larger homes that involve longer commutes, offsetting commuting time savings from WFH. The couple-level estimates corroborate the individual-level findings: WFH intensifies preexisting gender

[1] As of early 2025, 26% of work days are performed from home in the US, from 5% before the pandemic. In Germany, 25% of employed persons work from home at least partly (Alipour *et al.*, 2024, Fig. 1). These shares have stayed remarkably stable since 2022, despite recurring media attention to a return-to-office "trend" which has yet to bear out in the data.

gaps in caregiving and paid labor in the average couple. Three mechanisms help explain these patterns: individuals with more conservative gender norms use WFH to reinforce traditional divisions of labor; men's greater economic bargaining power allows them to protect time for paid work; and greater childcare demands–disproportionately borne by women–pull their time toward unpaid labor.

To date, we lack systematic evidence on how the WFH shock affects individuals' time use or couples' division of labor in the post-Covid economy. Importantly, earlier findings may not generalize as the scale, access, and attitudes toward WFH (by both firms and workers) have changed fundamentally since the pandemic (Barrero *et al.*, 2023). Harrington and Kahn (2023) find that pre-Covid, WFH appeared to reduce motherhood employment penalties in traditionally inflexible occupations such as finance. Similarly, pre-Covid data from Germany suggest that WFH correlates with increased paid hours, more so for women than men (Arntz *et al.*, 2022). Descriptive evidence from the pandemic suggests that WFH supported college-educated mothers' employment but also intensified their total workloads; meanwhile, household divisions of labor shifted only temporarily toward greater parity (Goldin, 2022; Lyttelton *et al.*, 2020). In the Netherlands, von Gaudecker *et al.* (2024) find that both parents in WFH-feasible jobs increased childcare time equally, narrowing the gender care gap as men more often held teleworkable jobs. Observational data from the US similarly indicate that mothers and fathers working remotely increased childcare time equally during the crisis (Pabilonia and Vernon, 2023). Yet importantly, pandemic-era conditions (especially school closures) likely inflated the burdens and understated WFH's longer-term potential, especially for women (Goldin, 2022). Schüller (2025) corroborates this argument by showing that by 2023, fathers' ability to WFH reduces the parental gap in domestic work, by increasing mothers' market hours and reducing their time in domestic work. However, WFH does not appear to affect the division of childcare among couples.

My main contribution is to provide *causal* evidence on WFH and the gender division of (un)paid labor in the post-Covid economy. I estimate the effects of *realized* WFH using an IV strategy that uses individual-level exposure to the WFH shock. Importantly, the panel structure allows me to substantiate identifying assumptions–e.g., by testing for parallel time-use trends prior to the WFH shock. A key novelty is to identify mechanisms, i.e., how WFH interacts with gender norms, bargaining power, and childcare needs. To this end, I expand on prior work by including non-parents, thereby accounting for gendered effects absent parenting constraints. The rich data also allow me to analyze cross-partner spillovers and directly assess how WFH affects within-couple gender gaps. Importantly, I account for time reallocation toward leisure–a key aspect of well-being that is typically lacking in the literature.

The core analysis uses a panel of 5,505 individuals employed in 2019 and reinterviewed in 2022. The SOEP includes annual information on time use across key activities (leisure, caregiving, domestic work, paid work) on an "average weekday" as well as hours worked from home in 2022.

Correlations between WFH and time use are likely confounded by unobserved factors driving self-selection into WFH. To address this endogeneity, I instrument realized WFH with a pre-pandemic measure of *WFH feasibility*, i.e., whether an individual's job *could* be done remotely in 2019. This measure strongly predicts realized WFH in 2022 and thus captures exposure to the Covid-induced WFH shock. The validity of the IV approach hinges on the exogeneity of WFH feasibility: First, because the pandemic was unanticipated, it is unlikely that individuals selected into jobs based on expected WFH opportunities.[2] Second, I predict WFH feasibility in the SOEP using a classification model trained on the 2018 BIBB-BAuA Employment Survey, based on pre-pandemic characteristics–occupation, industry, skill requirement, age, gender, and immigration background. This imputation purges potential bias from subjective influences (coming from e.g., employer policies or living conditions) and thus approximates whether a job can be done remotely solely based on its task profile. There are no gender differences in WFH feasibility, on average. Third, event-study results show no differential time-use trends by WFH feasibility before the pandemic, supporting the exogeneity assumption; namely, that WFH feasibility is uncorrelated with co-determinants of time-use changes.

The 2SLS estimates reveal pronounced gender asymmetries in time-use responses to WFH. Among women, an additional hour of WFH per weekday increases domestic work by 0.10 hours and leisure by 0.12 hours. Scaled to a full 7-hour WFH day, these translate to a 28% rise in domestic work and a 10% increase in leisure in 2022 relative to 2019. The impacts on care work and Do-it-yourself (DIY) activities are also positive but insignificant. These increases are mirrored by a drop in time spent on job-related activities (including commuting) by 0.27 hours per hour worked remotely, reflecting both commuting time savings and fewer paid hours. Women who WFH at least once per week reduce paid hours by 0.54 hours or 8%, on average. By contrast, WFH *reduces*

[2] Again, the key observation here is that the nature and perception of WFH arrangements have fundamentally changed after the pandemic such that even individuals with pre-Covid WFH experience should be considered "treated".

the likelihood of exiting employment relative to non-WFH peers, with effects nearly twice as large for women as for men. This aligns with pre-Covid evidence that WFH promotes labor force attachment for women (and mothers specifically) at the extensive margin (Harrington and Kahn, 2023). Thus crucially, the findings reveal a mixed pattern in women's labor supply: improved attachment at the extensive margin but reduced participation at the intensive margin.

Men, by contrast, show only small time-use responses to WFH. Domestic work rises modestly by 0.04 hours per hour of remote work, while effects on paid hours, leisure, and caregiving are insignificant. The muted response may seem puzzling given the time savings from reduced commuting. However, this can be reconciled with evidence that WFH prompts moves to larger homes with longer commutes, and lowers the likelihood of leaving employment, that is, eliminating commuting entirely.

At the couple level, IV results show that WFH amplifies gender gaps in paid and unpaid labor between heterosexual partners. Domestic and caregiving workloads shift disproportionately to women when only they work remotely, whereas effects are smaller and statistically weaker when only men do. When both partners WFH, gender gaps in paid hours and caregiving widen significantly. For the average couple, the gender gaps widen by 22% for caregiving and 18% for paid hours, relative to 2019. Gaps in leisure and DIY time remain unchanged, on average. Overall, the findings confirm that the transition to remote work reinforces rather than alleviates pre-existing gender disparities.

I identify three mechanisms that help explain the gendered effects of WFH. First, an intra-household bargaining channel: WFH slows wage growth for women relative to non-WFH peers, consistent with firms sharing part of WFH's amenity value with female employees, who tend to value flexibility more than men (Nagler *et al.*, 2024). I find no wage growth penalty for men. This asymmetry weakens women's bargaining position within the household, which tends to reinforce traditional divisions of labor. Consistent with this interpretation, the marginal effect of WFH on individuals' non-market labor declines with their household income share, for both men and women. Men, who hold higher bargaining power on average, are better able to shield time for paid work. Second, the effects of WFH interact with political leaning, a proxy for gender role attitudes: more conservative individuals, both women and men, use WFH to reinforce traditional divisions of labor. Third, childcare needs moderate the impact of WFH. Among women, the shift from paid labor to caregiving is concentrated in households with young children. For men with young children, caregiving time also rises with remote work, but the effect is significantly smaller and accompanied by an *increase* in paid hours. I do not find that WFH affects the presence (or the number of) children in the household, nor marital status.

The evidence challenges the idea that increasing job flexibility alone necessarily fosters gender equality (Goldin, 2014, 2022; Alon *et al.*, 2020). While WFH expands individual autonomy and lowers the cost of (temporal and locational) flexibility, its interaction with preexisting gender norms and unequal bargaining power, typically favoring men, tends to reinforce the traditional gender division of labor. Institutional features, such as Germany's joint taxation system ("Ehegattensplitting"), further incentivize using WFH arrangements to support main-earner households. Thus, lacking institutional and normative shifts, the Big Shift to WFH is more likely to entrench rather than mitigate gender inequality at home and in the labor market.

The paper proceeds with a description of the data and variables of interest in Section 2. Section 3 lays out the empirical strategy and the identifying assumptions. Section 4 presents the main results. Section 5 explores heterogeneity and potential mechanisms explaining the results. Section 6 concludes.

2 Data and variables of interest

2.1 German Socioeconomic Panel (SOEP)

The SOEP is a representative panel that currently surveys about 22,000 households in Germany annually. Version 39 includes data until 2022.

Time-use data and WFH status. The time-use questionnaire elicits hours spent on various activities on an "average weekday", including the following categories:[3]

1. *Work & education* (including commuting, and trainings)
2. *Care work* (including childcare and elder care)
3. *Domestic work* (including laundry, cooking, cleaning, and running errands)
4. *Do-it-yourself (DIY) activities* (repairs, gardening)
5. *Leisure* (including sleep, physical exercise, and hobbies)

Employed individuals also report their *weekly* working hours. For comparability with the time-use data, these entries are divided by five to obtain the average number of hours worked per weekday.

[3] I group some separate entries into one category; specifically, the entries for sleep, physical exercise, and "other leisure activities" are grouped into one "leisure" category. The results for "work" and "education" are combined into one group. I also count "running errands" as domestic work, and combine childcare and eldercare into one "care work" category.

The 2022 wave includes information on workers' WFH status. I use a dummy for *WFH at least once per week* and *average hours worked from home per weekday* (weekly hours divided by five) as the two explanatory variables of interest.

Individual-level panel. The main analysis uses a sample of individuals aged 16–64 and employed in 2019 (excluding military personnel, marginally employed, and freelancers), who completed the time-use questionnaire in 2019 and 2022. (2,513 men, 2,992 women). Table 1 presents summary statistics by gender. By 2022, 26% of women and 32% of men WFH at least once per week. Remote workers spend about two-thirds of their working hours from home, with no significant gender difference. Across all employed individuals, women spend 19% of their working hours from home and men 25%.[4]

The likelihood of holding a WFH feasible job in 2019 (described in the next section) does not differ significantly by gender. Unsurprisingly, time-use patterns reveal pronounced gender gaps. In 2019, women spent similar time on leisure (8.5 hours on an average weekday) but less on their job (including commuting and education) than men (7.4 vs. 9.3 hours). Women devoted one hour more than men to domestic work (2.6 vs. 1.5 hours) and more than twice as much to caregiving (1.9 vs. 0.8 hours). Women spent marginally less time on DIY activities (0.4 vs. 0.5 hours).

Couple-level panel. The data include 875 (heterosexual) cohabiting couples observed in 2019 and 2022, in which both partners are employed in 2019 and provided complete time-use data. The outcomes at the couple-level are gender gaps in time use–measured as the female-male difference in hours spent on an average weekday. Appendix Table B.1 reports the summary statistics. Consistent with the individual-level data, 2019 WFH feasibility does not differ within couples, on average; but men are more likely to work remotely than their partners by 2022.

2.2 BIBB-BAuA Employment Survey 2018

To obtain a binary measure of WFH feasibility at the job level, I use information from the 2018 wave of the BIBB-BAuA Employment Survey (Hall *et al.*, 2020).[5] The survey elicits working conditions and job features from a representative sample of the Ger-

[4] These WFH rates in 2022 align with estimates from other representative worker and company surveys, including the ifo Business Survey and the German Mikrozensus (Destatis).

[5] The survey is conducted by the Federal Institute for Vocational Education and Training (BIBB), and the Federal Institute for Occupational Safety and Health (BAuA).

Table 1: Summary statistics (individual level)

	(1) Women $N = 2992$	(2) Men $N = 2513$	(3) Difference (Women-Men)
A. WFH variables			
WFH at least 1x/week, 2022	0.258 (0.438)	0.322 (0.467)	-0.064 (0.012)***
Weekly hours WFH, 2022 (all)	5.341 (10.977)	8.349 (14.109)	-3.008 (0.338)***
Weekly hours WFH, 2022 (only WFH 1+/week)	19.642 (12.729)	24.824 (13.744)	-5.182 (0.667)***
Share of hours WFH, 2022 (all)	0.187 (0.467)	0.253 (0.765)	-0.066 (0.018)***
Share of hours WFH, 2022 (only WFH 1+/week)	0.606 (0.687)	0.683 (1.157)	-0.076 (0.048)
predicted WFH feasibility (0/1), 2019	0.508 (0.500)	0.504 (0.500)	0.004 (0.014)
predicted WFH feasibility [0,1], 2019	0.373 (0.291)	0.370 (0.298)	0.002 (0.008)
B. Main outcome variables			
B.1 *Time use (hours on average weekday)*			
Hours leisure (incl. exercise and sleeping), 2019	8.541 (1.792)	8.466 (1.838)	0.075 (0.049)
Δ Hours leisure (incl. exercise and sleeping)	0.156 (1.982)	0.317 (2.110)	-0.161 (0.055)***
Hours working (incl. studying and commuting), 2019	7.470 (2.630)	9.308 (1.922)	-1.838 (0.063)***
Δ Hours working (incl. studying and commuting)	-0.767 (3.388)	-1.060 (3.123)	0.293 (0.088)***
Hours care work, 2019	1.923 (3.458)	0.814 (1.532)	1.109 (0.074)***
Δ Hours care work	0.175 (3.924)	0.146 (1.859)	0.029 (0.085)
Hours domestic work (incl. errands), 2019	2.590 (1.347)	1.496 (1.041)	1.093 (0.033)***
Δ Hours domestic work (incl. errands)	0.293 (1.530)	0.462 (1.431)	-0.168 (0.040)***
Hours DIY activities, 2019	0.373 (0.624)	0.473 (0.720)	-0.100 (0.018)***
Δ Hours DIY activities	0.297 (0.837)	0.420 (1.051)	-0.122 (0.025)***
B.2 *Hours worked (weekly)*			
Weekly hours worked, 2019	32.027 (12.090)	41.219 (9.579)	-9.192 (0.298)***
Δ Weekly hours worked	-3.580 (14.258)	-4.850 (14.313)	1.270 (0.386)***
C. Further variables			
Age, 2019	45.423 (10.904)	45.397 (11.227)	0.027 (0.299)
Partner in HH (0/1), 2019	0.686 (0.464)	0.756 (0.429)	-0.070 (0.012)***
Children under 14 in HH (0/1), 2019	0.355 (0.478)	0.353 (0.478)	0.002 (0.013)
No of children u14 in HH, 2019	0.544 (0.840)	0.588 (0.934)	-0.044 (0.024)*
Commuting distance (km), 2019	14.098 (18.963)	18.218 (20.589)	-4.120 (0.554)***
Gross monthly wage (EUR), 2019	2489.632 (1629.863)	4061.217 (2609.004)	-1571.584 (57.724)***
Firm tenure in years, 2019	10.626 (10.472)	11.970 (10.931)	-1.344 (0.289)***
Firm with 100+ employees (0/1), 2019	0.623 (0.485)	0.686 (0.464)	-0.063 (0.013)***
Moved b/w 2019 and 2022 (0/1)	0.103 (0.304)	0.111 (0.314)	-0.007 (0.008)
No of rooms, 2019	4.376 (1.791)	4.380 (1.831)	-0.004 (0.049)
Log living space, 2019	4.634 (0.427)	4.632 (0.445)	0.002 (0.012)
Exited employment b/w 2019 and 2022 (0/1)	0.123 (0.328)	0.097 (0.296)	0.026 (0.008)***
Changed employer b/w 2019 and 2022 (0/1)	0.186 (0.389)	0.145 (0.352)	0.041 (0.011)***
Married (0/1), 2019	0.586 (0.493)	0.641 (0.480)	-0.055 (0.013)***
Single (0/1), 2019	0.252 (0.434)	0.275 (0.447)	-0.023 (0.012)*

Notes: The table reports variable means and standard deviations (in parentheses) by gender (not weighted). Column 3 reports gender differences and standard errors (in parentheses). The sample includes individuals employed and aged 16–64 in 2019, who are observed again in 2022. The operator Δ denotes (individual-level) changes between 2019 and 2022. The data are from the SOEP. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$

man workforce in 2018. I exclude marginal workers ("Minijobs"), military personnel, and freelancers, leaving a sample of 16,765 individuals.

Measuring WFH feasibility. When respondents in the BIBB-BAuA survey indicate that "WFH is impossible in my job" when asked whether they would accept a WFH offer, their job is coded as not WFH-feasible. In addition, I assume that a job cannot be worked from home if the respondent indicates *frequent* exposure to any of the following conditions:

1. Lifting or carrying loads of more than 10kg (women)/ 20kg (men)
2. Being exposed to smoke, dust, gases, or vapors
3. Being exposed to cold, heat, moisture, humidity, or droughts
4. Working with oil, grease, dirt
5. Handling microorganisms such as pathogens, bacteria, molds, or viruses
6. Working the majority of working hours outdoors
7. Repairing and renovating
8. Protecting, guarding, monitoring, regulating traffic
9. Cleaning, waste disposal, recycling
10. Monitoring, control of machines, plants, and technological processes

These additional restrictions help ensure that the measure captures jobs in which a substantial fraction of tasks can realistically be done from home. Under this definition, 39% of jobs are classified as WFH-feasible. This is broadly consistent with evidence that at the peak of the pandemic, about one-third of employed persons in Germany regularly worked from home (Alipour *et al.*, 2024, Fig. 1).

Predicting WFH feasibility in the SOEP. I capture individuals' exposure to the Covid-induced WFH shock by the *WFH feasibility* of their 2019 job. The variable is constructed in two steps: First, I train a supervised classifier model using data from the 2018 BIBB-BAuA Employment Survey to estimate (binary) WFH feasibility as a function of job and worker characteristics. Second, I apply the model to predict WFH feasibility for individuals in the SOEP sample.

One advantage is that this approach reduces the influence of idiosyncratic factors embedded in self-assessed WFH feasibility (e.g., due to employer policy or living conditions). As such, the resulting measure more accurately reflects exposure to the WFH shock derived from the task profile of the job. Second, the resulting variation is at the individual level (rather than the occupation level), allowing more precise identification.

The classification task is performed by a logistic model that estimates binary WFH feasibility as a function of occupation (KldB 3-digit, 139 categories), industry (NACE 2-digit, 85 categories), skill requirement (KldB 5-digit, 4 categories), age, age-squared, gender (0/1), and immigrant background (0/1). Including demographic characteristics can add power to the model by capturing sorting into certain jobs within occupations or industries. For identification purposes, I assume that this selection is orthogonal to anticipated WFH opportunities. This is plausible given that the Covid pandemic and the ensuing WFH shock were unexpected and fundamentally changed the nature (i.e., perception and organization) of WFH arrangements.

Predictors for the logistic model are selected by a LASSO regression with a penalty chosen by a 5-fold cross-validation to minimize the log-likelihood loss. The LASSO regularization penalizes coefficients, shrinking irrelevant predictors to zero and reducing overfitting. The cross-validation on random sub-samples improves the selection of the regularization parameter (lambda), which controls the strength of the penalty.[6]

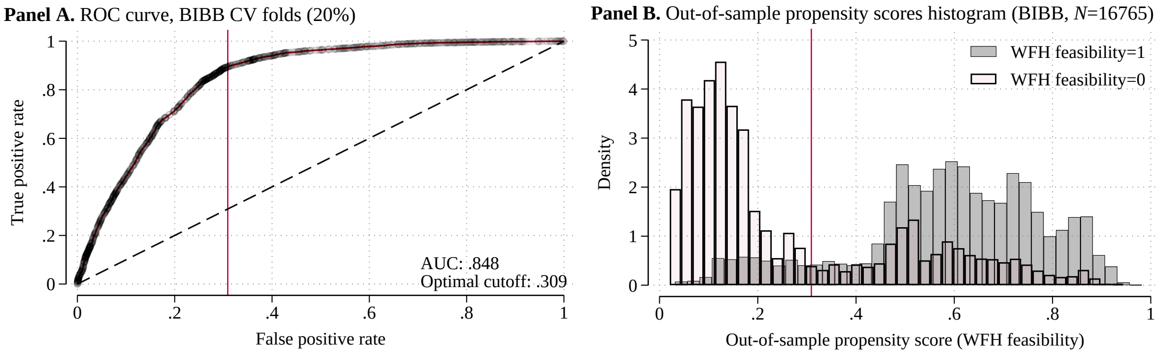
The LASSO procedure selects 209 out of 231 features. Model performance is summarized in Figure 1. Panel A depicts the ROC curve using out-of-sample predictions for WFH feasibility from the five cross-validation folds. The ROC curve visualizes the trade-off between true and false positives across various classification thresholds. The AUC (area under the curve) of 0.848 indicates an 84.8% probability that a randomly selected WFH-feasible job is assigned a higher propensity score than a randomly chosen non-feasible one.[7] By conventional standards, an AUC above 0.8 indicates strong predictive performance. The vertical line at 0.309 marks the classification cutoff that maximizes the difference between the true positive and false positive rate (Youden's J). This classification approach imposes equal weight to false negatives and false positives. Panel B plots the distribution of propensity scores by WFH feasibility status. At the Youden's J optimal cutoff (vertical line), the model correctly classifies 77% of cases (accuracy).

In the second step, I apply the trained model to predict WFH feasibility in the SOEP sample based on 2019 characteristics. This yields a continuous and a binary measure of (predicted) WFH feasibility, using the propensity scores and the classification threshold, respectively. There is no statistical gender difference in WFH feasibility (Ta-

[6] I use STATA's `cvlassologit` (available on SSC).

[7] A randomly guessing model would produce an AUC of 0.5, whereas a perfect model would produce an AUC of 1. The AUC summarizes model performance independently of any specific classification cutoff.

Figure 1: Out-of-sample performance of the binary prediction model



Notes: Panel A shows the ROC curve using out-of-sample predictions for WFH feasibility from the five cross-validation folds. Panel B plots histograms of out-of-sample propensity scores by WFH feasibility (0/1). The vertical lines at 0.309 highlight the classification cutoff that maximizes the difference between the true positive rate and false positive rate (Youden's J). Data are from the BIBB-BAuA Employment Survey 2018.

ble 1). The next section documents that WFH feasibility strongly predicts individuals' WFH status in 2022 and thus appropriately captures exposure to the Covid-induced WFH shock.

3 Empirical Strategy

3.1 Difference-in-differences (DiD) approach

At the individual level, I estimate a standard dynamic DiD specification, in which WFH feasibility in 2019 corresponds to (time-invariant) treatment status:

$$y_{it} = \sum_{k \neq 2019} [\beta^k \mathbb{1}(k=t) \times WFH\_FEAS_i] + \gamma_i + \gamma_t + \epsilon_{it}, \tag{1}$$

where y_{it} denotes an outcome (e.g., hours spent on domestic work) for individual i in year t. WFH_FEAS_i is WFH feasibility (binary (0/1) or continuous [0,1]) of i's job in 2019. The specification includes the standard individual (γ_i) and year γ_t fixed effects. The error term ϵ_{it} captures unobserved shocks, which are assumed to be uncorrelated with the treatment-year interactions. Then, the coefficients β^k trace the differential evolution of outcomes for individuals with WFH-feasible jobs relative to 2019 (the reference year). Standard errors are clustered by individual.

Parallel trends assumption. The validity of the DiD design rests on the parallel trends (PT) assumption. Given the unexpected nature of the Covid-induced WFH shock, the PT assumption is unlikely to be violated by selection on expected returns to WFH. A more salient concern is that WFH feasibility may be correlated with other determinants of time-use changes. For instance, WFH-feasible jobs are more common among highly educated workers and typically involve greater cognitive task content (see e.g., Alipour *et al.*, 2023). If time-use changes differ systematically across education groups, the DiD estimates may pick up a pre-existing trend rather than reflecting the causal effects of WFH. Checking for pretrends in the DiD plot offers a diagnostic for such confounding; though, their absence never guarantees that the PT assumption holds post-treatment.

3.2 Instrumental variable (IV) approach

The DiD estimates deliver intention-to-treat (ITT) effects due to non-compliance in the sense that individuals with a WFH feasible job in 2019 may not realize WFH opportunities. Realized WFH depends on factors such as employer policy, living arrangements, and individual preferences, each potentially influencing time use and thus limiting the ability of ITT estimates to measure the causal effect of *actually* working from home. To address this, I estimate an IV model in which realized WFH status in 2022 is instrumented by 2019 WFH feasibility. The IV estimates recover local average treatment effects (LATE) for compliers, i.e., individuals who work remotely *because* they were exposed to the WFH shock through their job's WFH feasibility.

Individual-level specification. I focus on long differences in time use between 2019 and 2022, as we are interested in the consequences of the WFH shock in the post-Covid economy. The individual outcome equation is given by

$$\Delta y_i = \beta_0 + \beta_1 \times WFH22_i + \epsilon_i, \tag{2}$$

with the corresponding first stage:

$$WFH22_i = \pi_0 + \pi_1 \times WFH\_FEAS_i + \vartheta_i, \tag{3}$$

where Δy_i is the *within-individual* change in time allocated to a given activity between 2019 and 2022. $WFH22_i$ is the individual's WFH status in 2022, corresponding to a dummy (working from home at least once per week) or a count variable (number of hours worked from home per week). WFH status is instrumented by (predicted) WFH feasibility (WFH_FEAS_i), measured continuously [0,1] or as binary (0/1). The model is estimated separately by gender $g \in \{f, m\}$.

Instrument relevance. For the IV strategy to be valid, we first require that the instrument be correlated with the endogenous variable (realized WFH). This is confirmed in Figure 2, which presents binned scatter plots of 2022 WFH status against WFH feasibility [0,1] by gender. Panel A uses a dummy for working remotely at least once per week in 2022; in Panel B, WFH status is the average number of hours worked remotely per weekday.

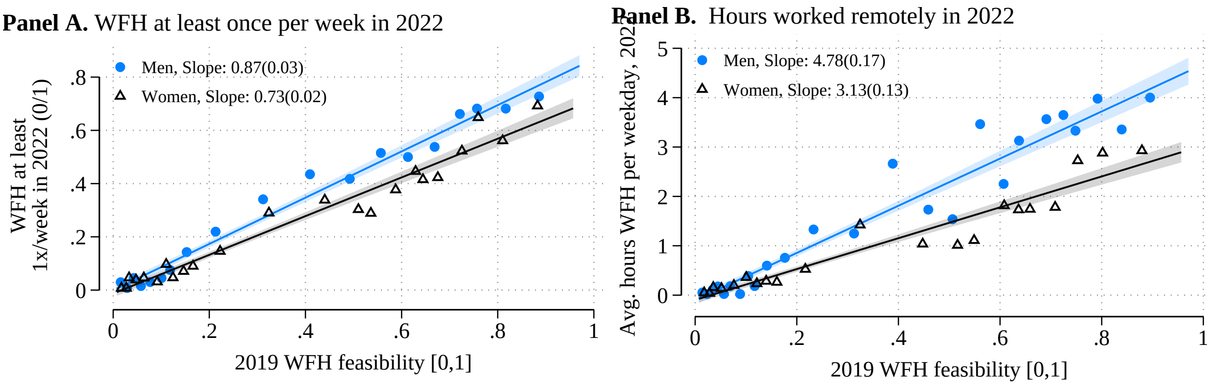
In both panels, the fitted lines show that WFH feasibility strongly predicts WFH status after the Covid shock. While WFH feasibility does not differ by gender on average, the first-stage slopes are flatter among women, suggesting they are less likely than men to realize their WFH potential. Moving from 0 to 1 in WFH feasibility raises the likelihood of working remotely at least once per week by 0.87 percentage points (p.p.) for men, compared to 0.73 p.p. for women. Similarly, WFH feasibility increases the average number of hours WFH per weekday by 4.78 among men, compared to 3.13 among women.[8]

These gender gaps in take-up are somewhat puzzling, given that women generally report a higher willingness to pay for WFH (Nagler *et al.*, 2024; Barrero *et al.*, 2021; Mas and Pallais, 2017). One explanation may be differential constraints or perceived returns: women may face tighter employer-imposed restrictions or anticipate lower benefits from WFH–e.g., in terms of productivity or promotion chances. Supporting this view, Emanuel *et al.* (2023) find that switching to WFH reduces feedback by senior colleagues more for women than men. In contrast, Bloom *et al.* (2024) run a randomized field experiment and find no gender differences in realized WFH, productivity, and promotion chances. Moreover, men's average commuting distance in 2019 is 4 kilometers longer than women's (Table 1), suggesting that time-savings benefits–and thus the incentive to WFH–may be stronger for men.

Exclusion restriction (ER). Second, the ER requires that WFH feasibility affect within-individual *changes* in time use only through realized WFH. Importantly, this means that any correlates of WFH feasibility are not problematic as long as they only influence time use levels. While the ER cannot be tested directly, three considerations support its plausibility: First, WFH feasibility is measured pre-Covid (2019) such that it is unaffected by the WFH shock. Second, individuals are unlikely to have selected into these jobs based on WFH opportunities created by the pandemic. Third, indi-

[8]These gender differences also hold conditional on total hours worked: Using the *fraction* of total hours worked from home as the dependent variable, the slopes are 0.73 among men, compared to 0.52 among women.

Figure 2: 2019 WFH feasibility and WFH status in 2022



Notes: The figure shows binned scatter plots of WFH status in 2022 against WFH feasibility [0,1], estimated based on 2019 characteristics by gender (see Section 2.2 for details). In Panel A, WFH status is a dummy for working remotely at least once per week in 2022; in Panel B, WFH status is the number of hours worked from home per week. Observations are grouped into equal-sized bins. Slope estimates (standard errors in parentheses) come from an OLS regression of WFH status on WFH feasibility. The sample includes individuals employed and aged 16–64 in 2019, who are observed again in 2022. Standard errors are clustered by individual. The data are from the SOEP.

rect tests of the ER deliver no evidence of differential pretrends by WFH feasibility, suggesting that the measure is not significantly correlated with co-determinants of time-use changes.

Limitation. A limitation of the IV strategy is that, by relying on a single instrument, it cannot separately identify the extensive margin (whether individuals WFH at all) from the intensive margin (how many hours they WFH). Clearly, moving from 0 to 1 hour of WFH differs from moving from 1 to 2. The IV estimate of *hours worked from home* should therefore be interpreted as an average local effect, combining adjustments at both margins.

Couple-level specification. Among cohabiting couples, each partner's WFH status can be linked to changes in their own time use, their partner's time use, and the within-couple gender gap in time use. For example, we can investigate whether the male partner's WFH status affects the female partner's paid hours or household labor, and vice versa.

The corresponding IV model thus involves two endogenous variables (both partners' realized WFH in 2022) and two instruments (both partners' WFH feasibility) and is given by

$$\Delta y_c = \gamma_0 + \gamma_1 \times WFH22_c^f + \gamma_2 \times WFH22_c^m + v_c, \tag{4}$$

with the first-stage equations:

$$WFH22_c^f = \pi_0 + \pi_1 \times WFH\_FEAS_c^f + \pi_2 \times WFH\_FEAS_c^m + \varepsilon_c, \tag{5}$$

$$WFH22_c^m = \rho_0 + \rho_1 \times WFH\_FEAS_c^m + \rho_2 \times WFH\_FEAS_c^f + u_c, \tag{6}$$

where y_c is a partner's time use or the *within-couple* time-use gap (female minus male) for a given activity. Δ denotes changes between 2019 and 2022. $WFH22_c^f$ and $WFH22_c^m$ are the WFH status in 2022 of the female and male partner in couple c. Each WFH status is instrumented by both partners' WFH feasibility in 2019 ($WFH\_FEAS_c^f$ and $WFH\_FEAS_c^m$). Then, $\hat{\gamma}_1$ and $\hat{\gamma}_2$ deliver the causal impact of female and male partner WFH on a partner's time use (or the couple's gender gap) compared to a couple in which neither partner works from home. The combined effect ($\hat{\gamma}_1 + \hat{\gamma}_2$) gives the impact of both partners WFH compared to a non-WFH couple.

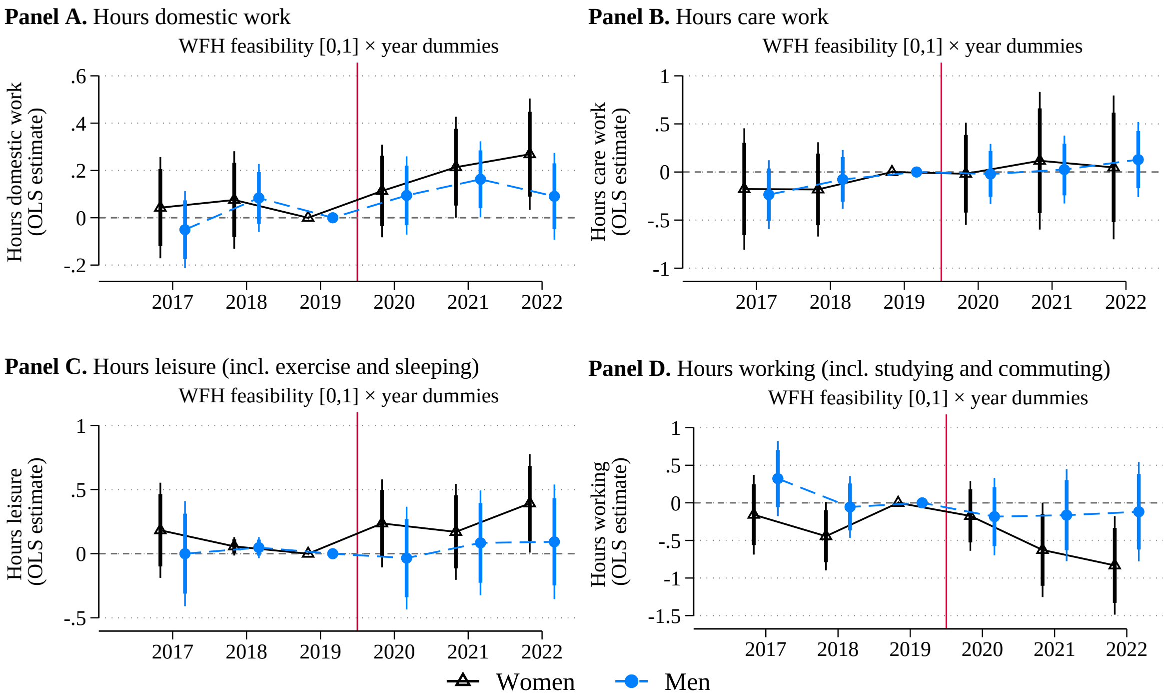
4 Main Results

Intention-to-treat (ITT) effects of WFH. Figure 3 presents DiD results for the four most time-intensive activities, separately for women and men.[9] The sample includes all individuals of the main sample. Appendix Figure A.1 plots DiD results for the subsample that is balanced between 2017 and 2022, yielding virtually identical results. All pre-treatment coefficients are insignificant at the 1% level. F-tests of joint insignificance fail to reject the null that pre-treatment coefficients are equal to zero at the 5% level in all but one case: for women's hours dedicated to work (including commuting), the pretrend is significant (p=0.039). However, the pretrend moves in the opposite direction of the treatment effect, suggesting that the treatment estimates may be conservative (Panel D). Overall, the evidence largely supports the PT assumption and suggests that WFH feasibility is not systematically correlated with determinants of time-use changes.

Following the WFH shock, women in WFH-feasible jobs shift time away from work (including commuting) toward domestic tasks and leisure. The impact estimates gradually intensify between 2020 and 2022. Among men, higher WFH feasibility is associated with a temporary increase in domestic work (Panel A) but does not lead to persistent changes in overall time allocation. For both genders, there is no discernible adjustment in caregiving hours, on average.

[9] Notice that there is no information on leisure in 2018; the individual coefficients are thus set to zero.

Figure 3: ITT effects of WFH on individual time use on an average weekday, balanced 2019–2022



Notes: The figure plots DiD estimates based on Equation 1. The sample includes individuals employed and aged 16–64 in 2019, balanced between 2019 and 2022. 99% and 95%-confidence intervals are drawn using standard errors clustered by individual. Data are from the SOEP.

Effects on individual time use. As explained, the DiD estimates deliver ITT effects, which are independent of realized WFH opportunities. Figure 4 reports the results from the IV model estimated by 2SLS, which accounts for compliance with treatment assignment. The estimates correspond to the local causal effects of WFH on 2019–2022 time-use changes for individuals who work remotely *because* their 2019 job could be done from home. Panel A uses the binary measure of WFH at least once per week as the explanatory variable; Panel B uses the average number of hours worked remotely per weekday. Panel C plots the 2019 mean of the dependent variables by gender for reference. Appendix Tables B.2 and B.3 present the corresponding point estimates. The first-stage F-statistics are above 500, confirming that WFH feasibility is a strong instrument.

Rows 1–5 present 2SLS estimates for separate outcomes of the time-use questionnaire. The last row shows the effect on the average number of paid hours worked per weekday.[10]

Among women, an additional hour of WFH per weekday significantly increases time spent on domestic tasks and leisure (including sleep) on an average weekday: 0.10 hours are reallocated to domestic work and 0.12 hours to leisure. Scaled to a full 7-hour remote workday, the effects translate to a 28% (10%) increase in domestic work and leisure relative to the 2019 means, respectively. Effects on DIY activities and care work are positive but insignificant. Women reduce time dedicated to their job (including education and commuting) by 0.27 hours or 26% per 7-hour WFH day. Of this, 0.10 hours reflect reduced paid work, suggesting that roughly two-thirds of the total reduction comes from commuting time savings and one-third from shorter working hours. Women who WFH at least once per week reduce paid working hours by 8%, on average.

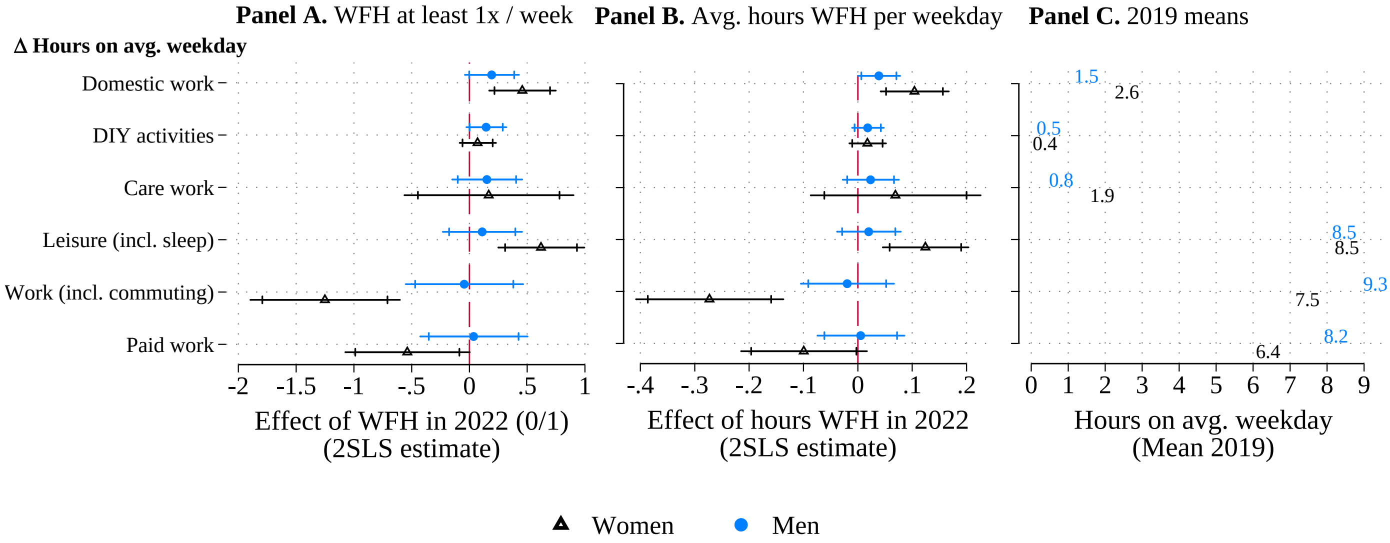
Among men, the only significant effect is a 0.04-hour increase in domestic work per additional hour WFH–equivalent to a 18% increase per 7-hour WFH day relative to 2019. This is smaller in absolute and relative terms than the corresponding effect for women. Effects on DIY activities, caregiving, and leisure are positive but statistically insignificant and weaker than for women. Interestingly, WFH has no discernible impact on men's time spent on work-related activities (including commuting) or paid hours worked. This result may seem counterintuitive, given that commuting-time savings are a central benefit of remote work. However, it can be reconciled with the findings (presented in Section 5) that WFH opportunities also influence individuals' work and living arrangements; in particular, by promoting moves to larger homes that involve *longer* commutes and by *reducing* the likelihood of exiting employment, i.e., not commuting at all.

Effects on couple-level gender gaps. The couple-level analysis allows us to study whether a partner's WFH status influences individuals' own time use. This setting also holds constant many contextual factors, including (re)location and household composition, as we focus on cohabiting couples.

Table 2 presents the IV estimates of WFH's impact on time use within couples. Panel A reports effects on the female partner's time use; Panel B shows results for the male

[10] *Weekly* hours worked are divided by five to match the weekday reference frame of the time-use module.

Figure 4: 2SLS estimates of the impact of WFH on individual time use



Notes: Panels A and B present 2SLS results based on Equations 2 and 3. Each coefficient comes from a separate IV regression of 2019–2022 time-use changes on WFH status, which corresponds to WFH at least once per week in 2022 (0/1) in Panel A and to the average number of hours WFH per weekday in 2022 (weekly hours WFH divided by five) in Panel B. The instrument is the WFH feasibility of individuals' 2019 job (see Section 2.2 for details). Panel C shows the 2019 means of the dependent variables by gender. The sample includes individuals employed and aged 16–64 in 2019, who are observed again in 2022. 95% and 90%-confidence intervals are drawn using standard errors clustered by individual. Data are from the SOEP. Appendix Tables B.2 and B.3 present the corresponding point estimates and first-stage diagnostics.

partner. Panel C relates each partner's WFH status to the within-couple gender gap in time use (female minus male hours per weekday). By construction, these estimates equal the difference between the coefficients in Panels A and B.

WFH status is defined as WFH at least once per week (0/1). Results using average remote hours per weekday are similar and reported in Appendix Table B.5. In both cases, the Lewis-Mertens statistics (g_{min}) exceed the critical values at a significance level of $\alpha = 0.05$ and a relative bias threshold of $\tau = 0.1$, rejecting the null of weak instruments.[11]

Panel A shows that the estimates mirror the individual-level results: Women's own WFH raises their time in domestic labor and lowers their time in market work (including commuting). The impact on care work is large (about 50% of the 2019 mean) but estimated imprecisely. In contrast, male partner WFH has no significant effect

[11] Lewis *et al.* (2024) generalize the Olea and Pflueger (2013) robust effective F-statistic for testing weak instruments based on a worst-case benchmark for the 2SLS bias to models with multiple endogenous regressors under heteroskedasticity and serial correlation.

on women's time use, on average. When both partners WFH, caregiving time rises significantly by 1.9 hours, while effects on leisure remain small and insignificant.

Panel B suggests that men's own WFH raises their domestic and care work, though the effects are much smaller than for women. When the female partner works remotely, men tend to reduce their participation in domestic and care work. Male WFH is also associated with more time spent on paid work, though the estimates are imprecise.

Taken together, the estimates indicate that WFH widens gender gaps within couples (Panel C). When only the woman WFH, the domestic work gap increases by 0.83 hours; when only the man WFH, it narrows, but to a lesser extent. When both partners WFH, the gap increases modestly by 0.11 hours per day (column 1). Scaling the coefficients by the average WFH rates of both partners, the impact is nearly zero.

A similar pattern emerges for caregiving. Male WFH does not significantly affect the care gap, but female WFH increases it. When both partners work remotely, the gender gap in care work rises by 1.2 hours. In the average couple, the gap increases by 0.34 hours or 23% relative to 2019 (column 2). DIY and leisure gaps show no significant change.

For paid work, both partners' WFH contributes to a widening gender gap (columns 5 and 6). While individual effects are not or only marginally significant, their sum is sizable: when both partners WFH, the paid hours gap increases by 1.4 hours per weekday (column 6). For the average couple, this implies a rise of 0.42 hours, or 18% relative to 2019.

In sum, the couple-level analysis corroborates the individual-level findings: among the population employed shortly before the Covid crisis, the Big Shift to WFH intensifies gender inequality in both paid and unpaid labor.

Table 2: 2SLS estimates of the impact of WFH (at least once per week) on couples' gender time-use gaps

	(1)	(2)	(3)	(4)	(5)	(6)
	Δ Time use for					
	Dom. work (incl. errands)	DIY activities	Care work	Leisure (incl. sleep, exercise)	Work (incl. trainings, commuting)	Paid work
Panel A: Effect on female partner's time use						
Female partner's WFH, 2022 (0/1)	0.647**	-0.182	1.311	-0.049	-1.209**	-0.518
	(0.293)	(0.155)	(0.893)	(0.328)	(0.610)	(0.517)
Male partner's WFH, 2022 (0/1)	-0.268	0.142	0.611	-0.054	0.108	-0.080
	(0.234)	(0.131)	(0.755)	(0.288)	(0.518)	(0.429)
Combined effect ($\hat{\gamma}_1 + \hat{\gamma}_2$)	.379	-.041	1.921***	-.103	-1.101**	-.597
	(.236)	(.117)	(.706)	(.273)	(.512)	(.436)
2019 DV mean	2.825	.401	2.571	8.498	7.019	6.063
Panel B: Effect on male partner's time use						
Female partner's WFH, 2022 (0/1)	-0.178	-0.256	-0.056	-0.055	0.096	0.279
	(0.267)	(0.169)	(0.375)	(0.395)	(0.536)	(0.491)
Male partner's WFH, 2022 (0/1)	0.449**	0.266*	0.755**	-0.084	0.067	0.505
	(0.198)	(0.143)	(0.312)	(0.326)	(0.451)	(0.419)
Combined effect ($\hat{\gamma}_1 + \hat{\gamma}_2$)	.271	.01	.699**	-.14	.163	.784*
	(.276)	(.165)	(.303)	(.361)	(.479)	(.451)
2019 DV mean	1.437	.582	1.057	8.295	9.352	8.364
Panel C: Effect on couple's time-use gap (W-M)						
Female partner's WFH, 2022 (0/1)	0.825**	0.074	1.367*	0.006	-1.305*	-0.797
	(0.381)	(0.218)	(0.786)	(0.458)	(0.747)	(0.636)
Male partner's WFH, 2022 (0/1)	-0.717**	-0.124	-0.144	0.030	0.041	-0.585
	(0.300)	(0.186)	(0.669)	(0.385)	(0.643)	(0.530)
Combined effect ($\hat{\gamma}_1 + \hat{\gamma}_2$)	.109	-.05	1.222**	.036	-1.264*	-1.381**
	(.35)	(.181)	(.607)	(.396)	(.648)	(.571)
Average effect ($\hat{\gamma}_1 \times \overline{WFH22}^f + \hat{\gamma}_2 \times \overline{WFH22}^m$)	.002	-.019	.341*	.011	-.357*	-.415**
	(.104)	(.054)	(.18)	(.118)	(.194)	(.17)
2019 DV mean	1.389	-.181	1.514	.203	-2.333	-2.301
#Couples	875	875	875	875	875	875
First-stage diagnostics						
Lewis-Mertens statistic (g_{min})	80.51	80.51	80.51	80.51	80.51	80.51
g_{min} critical values ($\alpha = 0.05, \tau = 0.1$)	20.22	20.22	20.22	20.22	20.22	20.22

Notes: The table presents 2SLS results based on Equations 4–6. Time-use gaps (Panel C) refer to the female-male differences in hours spent on a given activity on an average weekday within a couple. The instruments are the WFH feasibility [0,1] of each partner's 2019 job (see Section 2.2 for details). The sample includes couples in which both partners are employed and aged 16–64 in 2019, and observed again in 2022. Standard errors are clustered by couple and reported in parentheses. Data are from the SOEP. $^{***}p < 0.01, ^{**}p < 0.05, ^{*}p < 0.1$

5 Heterogeneity and Mechanisms

5.1 Effects on individual work and living conditions

To examine potential mechanisms, Table 3 reports 2SLS results of the impact of WFH on work and living conditions by gender. WFH status corresponds to a dummy for WFH at least once per week in 2022. Appendix Table B.4 presents results using average hours WFH per weekday.

For men, WFH increases the probability of moving by 6.5 p.p., with no significant effect for women (column 1). This gender difference is consistent with persistently higher job-related mobility among men (Jayachandran *et al.*, 2024). Conditional on moving, WFH changes housing conditions for both genders: the number of rooms increases significantly, and living space tends to expand relative to 2019 (columns 2 and 3); the effects are stronger with more intensive WFH (Table B.4). These findings substantiate evidence from the urban and real estate literature that WFH raises demand for space and facilitates moves to peripheral areas of cities (Akan *et al.*, 2025; Althoff *et al.*, 2022; Gupta *et al.*, 2022; Delventhal and Parkhomenko, 2023; Coskun *et al.*, 2024).[12]

WFH also reduces the likelihood of exiting employment (column 4).[13] The impact is larger for women, aligning with pre-Covid evidence that WFH narrows the motherhood employment penalty and improves women's labor market attachment (Harrington and Kahn, 2023). Importantly, the results highlight an asymmetry in female labor supply: WFH strengthens participation at the extensive margin, but reduces it at the intensive margin, as employed women cut back on hours.

Column 5 shows that WFH reduces the likelihood of changing employer, especially for women. This corroborates evidence from field and discrete choice experiments that WFH increases worker retention by conferring a valuable amenity, which is appreciated more strongly by women (Bloom *et al.*, 2024, 2015; Nagler *et al.*, 2024; Mas and Pallais, 2017). The reduction in women's paid hours when WFH is accompanied by a slower wage growth of about 0.20 log points (column 6). However, roughly 45% of this decline reflects a penalty on hourly wage growth of 0.09 log points (column 7). This is consistent with employers sharing part of the amenity value of WFH with female employees. Notably, no wage growth penalty is observed for men working remotely. This asymmetry suggests that WFH tends to weaken women's bargaining position in

[12] Unfortunately, the 2022 wave of the SOEP does not elicit commuting distance.

[13] Individuals who exit employment (by leaving the labor force or going into unemployment) have zero working hours (from home).

the household and potentially their ability to negotiate a more equal division of labor. Section 5.2 further investigates the intra-household bargaining channel.

Finally, WFH has no significant impact on marital status or the presence of children under 14 in the household (columns 8–10). Likewise, the number of children is not affected by WFH (not shown in the table), suggesting little short-run effect on household structure or fertility, on average.

Table 3: 2SLS estimates of the impact of WFH (at least once per week) on work and living conditions

	(1) Moved (0/1)	(2) #Rooms	(3) Log living space	(4) Left employm. (0/1)	(5) Employer change (0/1)	(6) Log gross labor income	(7) Log gross hrly wage	(8) Married (0/1)	(9) Single (0/1)	(10) Children <14 (0/1)
Panel A: Women										
WFH at least 1x/week in 2022	-0.017	0.301**	0.039	-0.076**	-0.160***	-0.199***	-0.091**	0.000	-0.027	0.027
	(0.029)	(0.118)	(0.026)	(0.031)	(0.036)	(0.050)	(0.046)	(0.026)	(0.017)	(0.036)
#Individuals	2,992	2,972	2,992	2,992	2,624	2,624	2,607	2,991	2,991	2,992
DV mean	.103	4.376	4.634	.123	.186	7.575	2.748	.586	.252	.355
First-stage coef.	0.39	0.39	0.39	0.39	0.43	0.43	0.43	0.39	0.39	0.39
	[738.6]	[735.3]	[738.6]	[738.6]	[784.3]	[772.2]	[777.9]	[739.2]	[739.2]	[738.6]
Panel B: Men										
WFH at least 1x/week in 2022	0.065**	0.230**	0.041*	-0.048**	-0.037	-0.005	0.043	0.010	-0.015	0.001
	(0.026)	(0.112)	(0.022)	(0.024)	(0.028)	(0.036)	(0.039)	(0.021)	(0.016)	(0.030)
#Individuals	2,513	2,488	2,513	2,513	2,270	2,269	2,256	2,511	2,511	2,513
DV mean	.111	4.38	4.632	.097	.145	8.122	2.979	.641	.275	.353
First-stage coef.	0.48	0.48	0.48	0.48	0.53	0.52	0.53	0.48	0.48	0.48
	[926.3]	[914.2]	[926.3]	[926.3]	[1001.0]	[985.7]	[995.8]	[927.0]	[927.0]	[926.3]
p-values test of equal effects	0.0337	0.6613	0.9467	0.4766	0.0065	0.0015	0.0258	0.7517	0.6090	0.5831

Notes: The table presents 2SLS results based on Equations 2 and 3, estimated separately by gender. The instrument is the WFH feasibility (0/1) of individuals' 2019 job (see Section 2.2 for details). The sample includes individuals employed and aged 16–64 in 2019, who are observed again in 2022. Columns 5–7 exclude individuals who are not employed in 2022. Kleibergen-Paap F-statistic for the first-stage regression in brackets. Standard errors are clustered by individual and reported in parentheses. Data are from SOEP. $^{***}p < 0.01,^{**}p < 0.05,^{*}p < 0.1$

5.2 Interactions with gender norms, bargaining power, and caregiving needs

To understand why WFH tends to reinforce a more traditional gender division of labor on average, I explore heterogeneity along three determinants of labor supply: norms, constraints, and bargaining. First, political leaning plausibly captures variation in gender norms, which influence preferences over the allocation of (un)paid labor. Second, the presence and age of children capture time constraints and caregiving demands, and condition how individuals can use the flexibility afforded by WFH. And third, individuals' relative income share within the household reflect intra-couple bargaining power and may influence time-use negotiations in the context of WFH.

I estimate separate 2SLS models with an interaction term between WFH and each dimension and report the marginal effects of WFH across their distributions; specifically the outcome equation at the individual level corresponds to:

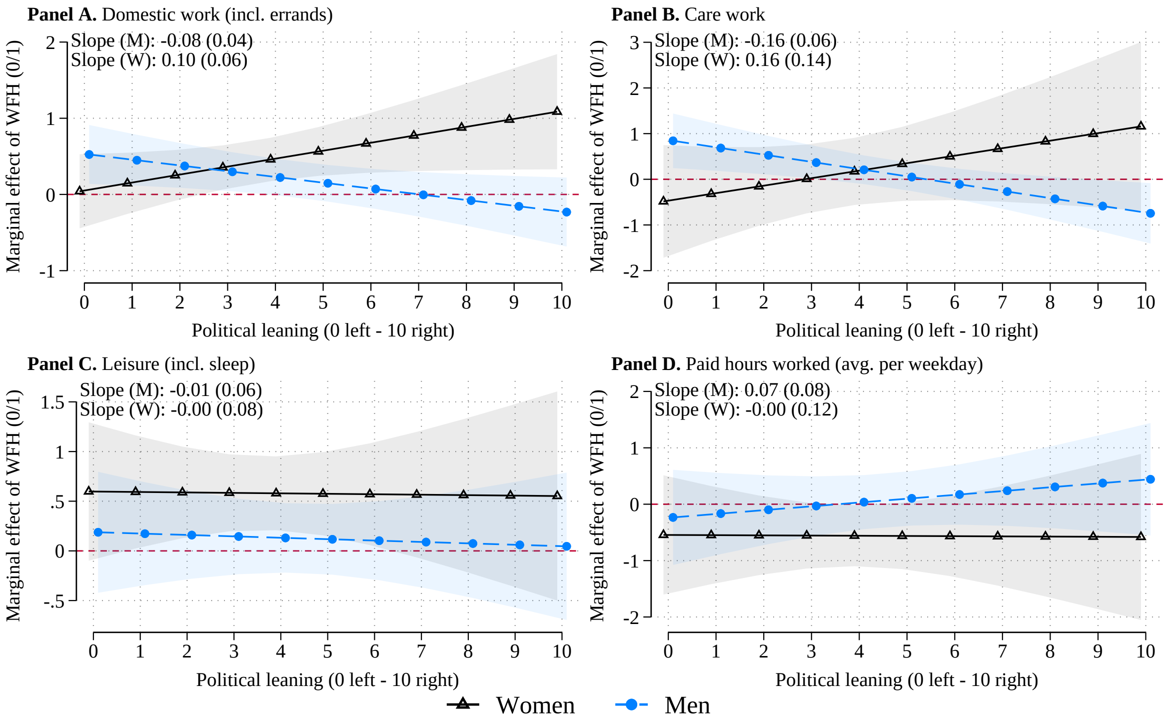
$$\Delta y_i = \alpha_0 + \alpha_1 \times WFH22_i \times Z_i + \alpha_2 \times WFH22_i + \alpha_3 \times Z_i + \epsilon_i, \tag{7}$$

where Z_i denotes a (time-invariant) candidate moderator. The 2SLS estimate is obtained by instrumenting WFH status (WFH at least once per week 0/1) and the interaction term with Z_i by WFH feasibility and an analogous interaction with WFH feasibility.

Gender norms. First, I examine effect heterogeneity by individuals' political leaning, which is elicited on a 10-point scale from 0 (left) to 10 (right). Political leaning typically correlates strongly with gender role attitudes (Nennstiel and Hudde, 2025; Oswald and Powdthavee, 2010). On average, men are somewhat more conservative than women (4.61 vs. 4.36). The question is whether individuals may use WFH differently in ways that reflect their gender norms. If so, we would expect WFH to reinforce traditional divisions of labor among more conservative individuals and to shift time-use toward a more egalitarian division among those with progressive attitudes.

Figure 5 plots the marginal effects of WFH at least once per week (0/1) on time-use changes across the distribution of individuals' political leaning. Panels A and B clearly support the gender norm channel: among women, the effects of WFH on domestic work and caregiving become more positive with increasingly conservative views. The pattern is reversed among men: WFH increases participation in household labor as attitudes become more progressive. Given average political attitudes, WFH interacts with gender norms to widen the gender gap in household labor. I find no clear interaction between WFH and political leaning for leisure and a suggestive interaction for paid hours: the marginal effect of WFH on working hours increases slightly with conservative views among men.

Figure 5: Marginal effects of WFH (0/1) on time use by political leaning (2SLS estimates)



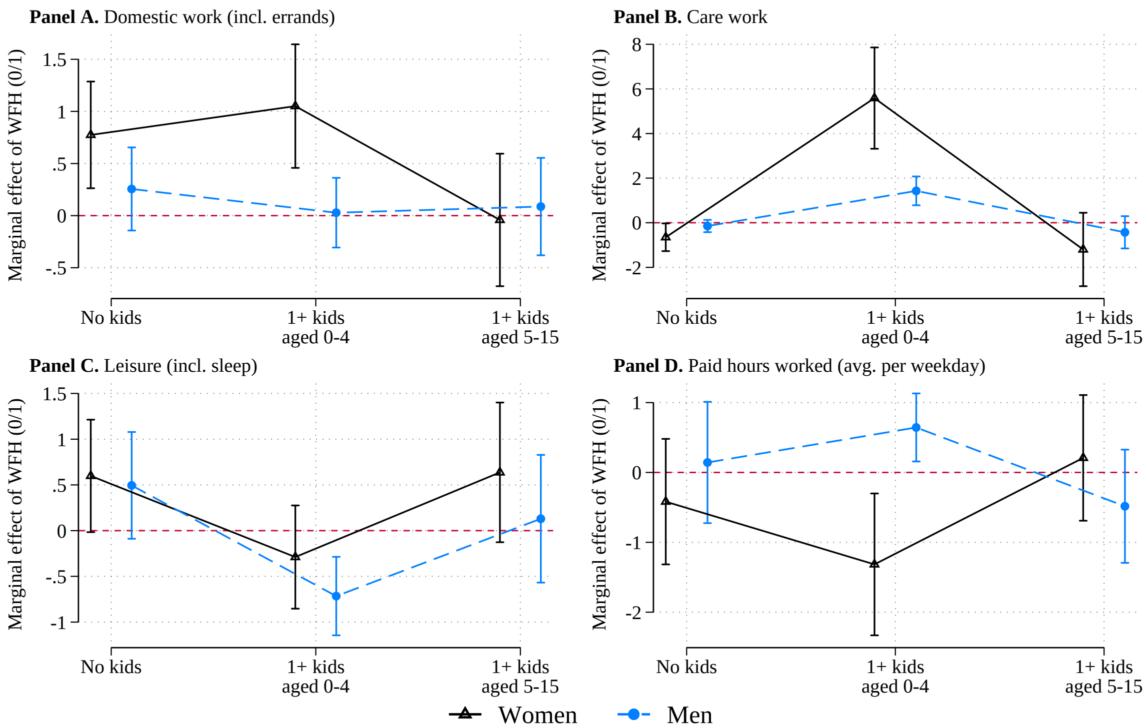
Notes: Panels A–D show 2SLS estimates based on Equation 7, estimated separately by gender. Plotted are the marginal effects of WFH at least once per week in 2022 (0/1) one time-use changes (2019–2022) across the distribution of individuals' political leaning, elicited on a 10-point scale from 0 (left) to 10 (right). The sample includes individuals employed and aged 16–64 in 2019. Confidence bands are drawn at the 95% level using standard errors clustered by individual. Data are from the SOEP.

Children. Second, childcare needs raise time constraints and the opportunity costs of alternative time uses, including leisure or market labor (Aguiar and Hurst, 2007). Consequently, relaxing the time budget constraint via WFH likely causes different reallocation of time depending on the presence and age of children. Primarily, parents are expected to divert more time toward household labor. I examine possible heterogeneity by caregiving needs using a categorical variable that distinguishes between households without children, those with at least one child aged 0–4, and those with at least one child aged 5–15.

Figure 6 plots the marginal effects of WFH at least once per week (0/1) on time-use changes by the presence and age of children. The results show clear gender asymmetries. Among women, WFH increases time spent on domestic and care work, particularly when a young child is present, consistent with heightened caregiving demands

(Panels A and B). These effects are muted or absent for older children or when no children are present. Among men, time use remains largely unaffected across child age groups. Care work increases slightly with young children, but the effect is small relative to that for women (Panel B). For leisure, WFH reduces time use when caregiving needs are high and increases time use for childless households. For paid work, WFH reduces hours primarily for women with young children. Interestingly, WFH has the opposite effect for men with young children. When caregiving demands are minimal or absent, WFH does not significantly affect paid work hours for either gender. Taken together, the results suggest that childcare needs amplify gendered responses to WFH as women reallocate time toward unpaid labor, while men's time allocation remains largely stable or shifts toward market work.

Figure 6: Marginal effects of WFH (0/1) on time use by presence of children (2SLS estimates)



Notes: Panels A–D show 2SLS estimates based on Equation 7, estimated separately by gender. Plotted are the marginal effects of WFH at least once per week in 2022 (0/1) one time-use changes (2019–2022) by the presence of children in the household, measured by 3 mutually exclusive categories (no children, at least one child aged 0-4, at least one child aged 5–15). The sample includes individuals with a cohabiting partner, and who are employed and aged 16–64 in 2019. Confidence intervals are drawn at the 95% level using standard errors clustered by individual. Data are from the SOEP.

Bargaining power. Third, I examine effect heterogeneity by individuals' labor share of household income among individuals with a cohabiting partner. A higher share approximates a greater economic bargaining weight in the collective household bargaining framework. The idea is that individuals with stronger bargaining positions may more effectively use WFH to negotiate favorable changes to the division of labor according to their preferences.

Figure 7 plots the marginal effects of WFH at least once per week (0/1) on time-use changes across the distribution of household income shares. In Panels A and B, the marginal effect curves are downward sloping for both genders, suggesting that WFH is used to reduce unpaid labor when bargaining power is higher. No clear effect is detectable for leisure. If anything, women with higher relative income use WFH to reduce leisure time and increase their market labor supply (Panel D). The opposite pattern holds for men, but the estimates are imprecise. Overall, these results align with bargaining models: partners with greater relative income–especially women–appear more able to shield their time for paid work and reduce the burden of unpaid labor under WFH. Ultimately, the bargaining channel strengthens the inequality in the division of labor as household income is disproportionately concentrated among men relative to women (63.22% vs. 36.25%).

5.3 Complier analysis

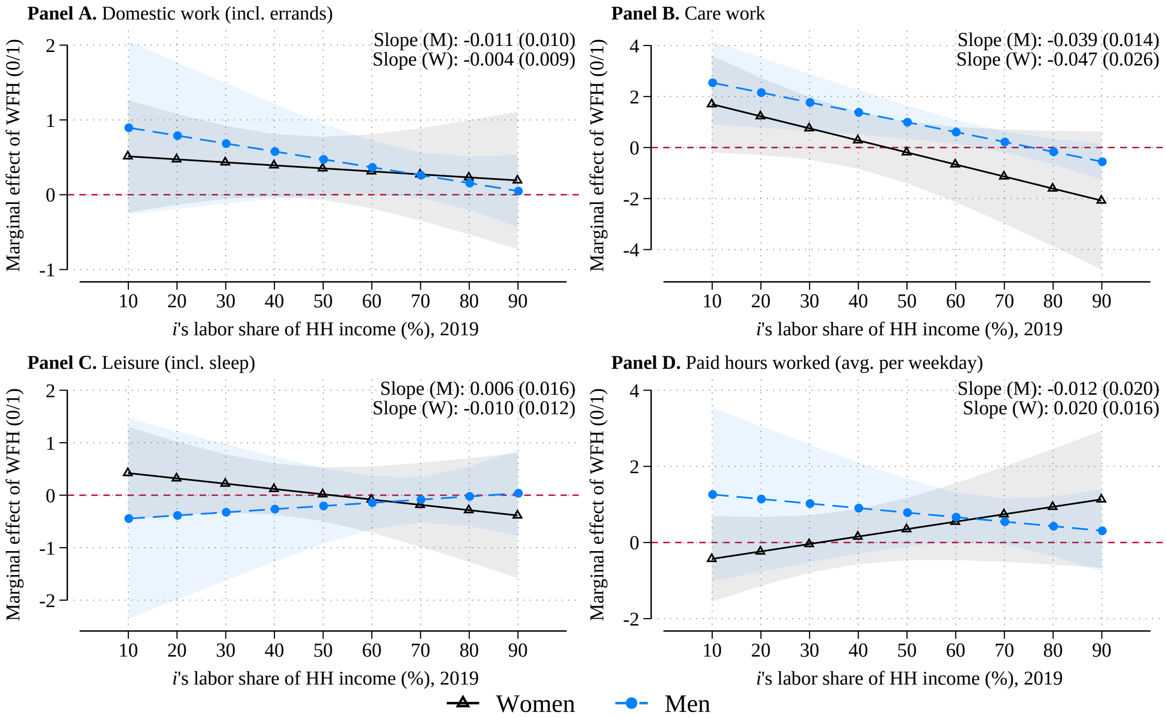
The IV estimates identify local average treatment effects (LATE) for the subpopulation of compliers, i.e., individuals who work remotely if and only if their 2019 job is WFH feasible. Results may not be generalizable if treatment effects are heterogeneous and compliers differ strongly from the broader sample. I assess differences in pre-determined characteristics using the methodology by Marbach and Hangartner (2020); Hangartner *et al.* (2021). The approach provides a simple way of identifying and characterizing compliers relative to the whole sample.[14]

Figure 8 presents covariate means with 95%-confidence intervals for the full sample and the complier subpopulation by gender. Appendix Table B.6 provides the corresponding point estimates, tests for mean equality, and standardized differences.

Overall, compliers resemble the broader sample along most observable dimensions. They are, on average, slightly younger (by less than two years), which is reflected in shorter firm tenures and lower monthly wages in 2019 (Panels G and H). Male compliers commute about 2 kilometers longer in 2019 than the male sample; no such

[14] Under the standard IV assumptions, observable always-takers (individuals assigned to the control condition who take the treatment) and never-takers (treated who refuse treatment) have identical covariate distributions as their non-observable counterparts, allowing direct estimation of their covariate means. Complier means are identified indirectly by subtracting the weighted covariate means of always-takers and never-takers from the overall sample mean (Marbach and Hangartner, 2020).

Figure 7: Marginal effects of WFH (0/1) on time use by labor share of household income (2SLS estimates)

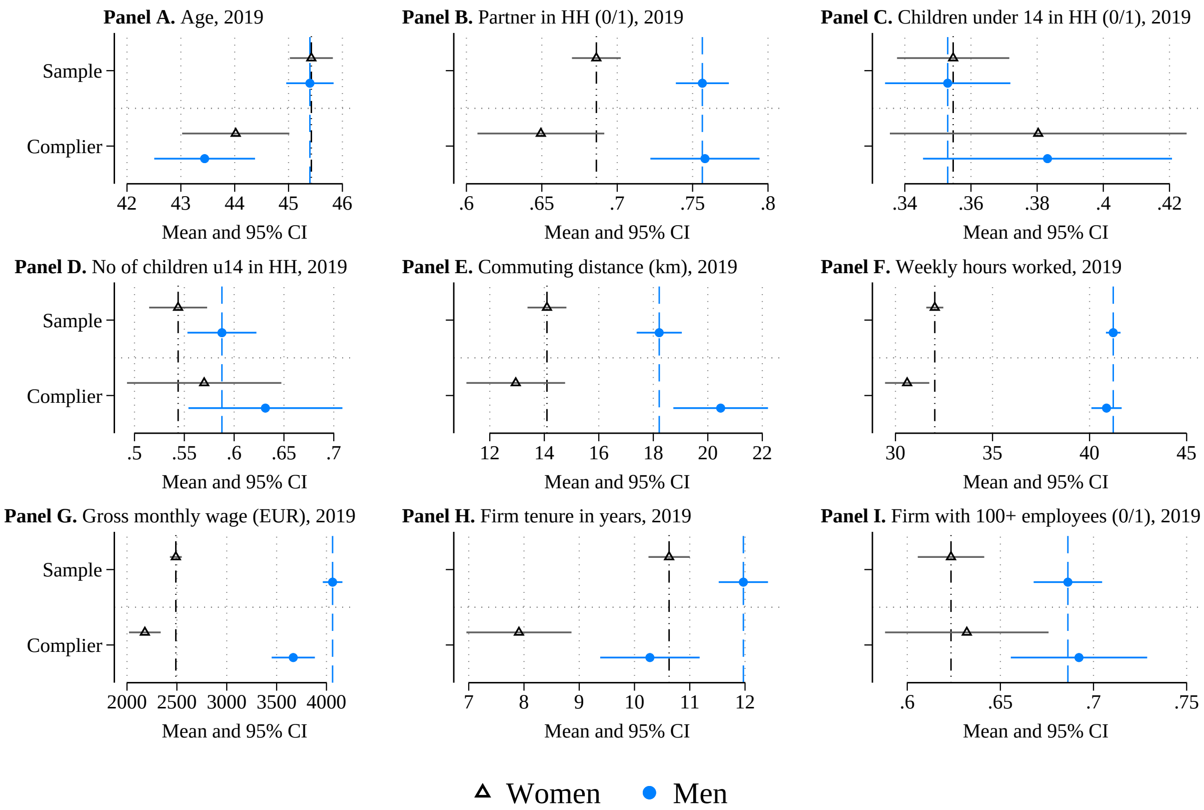


Notes: Panels A–D show 2SLS estimates based on Equation 7, estimated separately by gender. Plotted are the marginal effects of WFH at least once per week in 2022 (0/1) one time-use changes (2019–2022) across the distribution of individuals' labor share of household income. The sample includes individuals with a cohabiting partner, and who are employed and aged 16–64 in 2019. Confidence bands are drawn at the 95% level using standard errors clustered by individual. Data are from the SOEP.

difference appears for women. Among women, compliers work 1.4 fewer hours per week in 2019, compared to the sample average of 32 hours. Differences in family structure (having children under 14 or a partner in the household) and firm size are small and statistically insignificant.

Standardized differences fall below the conventional 0.25 threshold for considerable group imbalances for all covariates, except for firm tenure among female compliers, which is 2.7 years shorter than the sample (standardized difference = 0.29) (Imbens and Rubin, 2015). Taken together, the complier population is broadly similar to the overall sample, bolstering the interpretability of the LATE in this context.

Figure 8: Complier analysis



Notes: The figure plots covariate means for the whole sample and the complier subpopulation based on the methodology by Marbach and Hangartner (2020); Hangartner *et al.* (2021). The instrument is WFH feasibility (0/1), and the treatment is WFH at least once per week in 2022 (0/1). The sample includes individuals employed and aged 16–64 in 2019, who are observed again in 2022. 95%-confidence intervals are calculated using bootstrapped standard errors (1,000 replications). Appendix Table B.6 reports the corresponding point estimates and tests of equal means.

6 Conclusion

This paper investigates how the large-scale transition to WFH, triggered by the Covid-19 pandemic, impacts the gender division of labor. Drawing on plausibly exogenous individual-level variation in the exposure to the WFH shock, I estimate the effects of WFH on changes in time allocation to unpaid labor, paid work, and leisure. The results reveal marked asymmetries: women transitioning to WFH reduce time spent on paid work and shift it toward domestic labor and leisure. Men, by contrast, show little adjustment across time-use categories due to WFH, which can be explained by residential moves to larger, more distant homes that offset commuting time savings. Couple-level analyses of partners' time-use gaps bolster the finding that WFH exacerbates gender disparities. For the average couple, the gender gaps in market labor and unpaid work, notably in caregiving, intensify.

Three mechanisms help explain these patterns. First, individuals with more conservative gender norms use WFH to reinforce traditional divisions of labor. Second, men's greater economic bargaining power allows them to protect time for paid work. And third, greater childcare demands–disproportionately met by women–draw their time toward unpaid labor. These results challenge the view that greater job flexibility inherently fosters gender equality. When the autonomy afforded by WFH interacts with existing social expectations and institutional incentives (e.g., the tax system), it ultimately tends to reinforce rather than mitigate the gendered division of labor.

References

Aguiar, M. and Hurst, E. (2007). 'Life-Cycle Prices and Production', *American Economic Review*, vol. 97(5), pp. 1533–1559, ISSN 0002-8282, doi:10.1257/aer.97.5.1533.

Akan, M., Barrero, J.M., Bloom, N., Bowen, T., Buckman, S.R., Davis, S.J. and Kim, H. (2025). 'The New Geography of Labor Markets', NBER Working Paper 33582, National Bureau of Economic Research, doi:10.3386/w33582.

Aksoy, C.G., Barrero, J.M., Bloom, N., Davis, S.J., Dolls, M. and Zarate, P. (2022). 'Working From Home Around the World', *Brookings Papers on Economic Activity*, vol. Fall 2022, pp. 281–330, doi:10.1353/eca.2022.a901274.

Aksoy, C.G., Barrero, J.M., Bloom, N., Davis, S.J., Dolls, M. and Zarate, P. (2025). 'Working from Home in 2025: Five Key Facts', *EconPol Policy Brief*, vol. 9(73).

Alipour, J.V., Falck, O., Krause, S., Krolage, C. and Wichert, S. (2024). 'Working from Home and Consumption in Cities', CESifo Working Paper.

Alipour, J.V., Falck, O. and Schüller, S. (2023). 'Germany's capacity to work from home', *European Economic Review*, vol. 151, p. 104354, doi:10.1016/j.euroecorev.2022.104354.

Alon, T., Doepke, M., Olmstead-Rumsey, J. and Tertilt, M. (2020). 'This Time It's Different: The Role of Women's Employment in a Pandemic Recession', NBER Working Paper 27660, National Bureau of Economic Research, doi:10.3386/w27660.

Althoff, L., Eckert, F., Ganapati, S. and Walsh, C. (2022). 'The Geography of Remote Work', *Regional Science and Urban Economics*, vol. 93, p. 103770, ISSN 0166-0462, doi:10.1016/j.regsciurbeco.2022.103770.

Arntz, M., Ben Yahmed, S. and Berlingieri, F. (2022). 'Working from home, hours worked and wages: Heterogeneity by gender and parenthood', *Labour Economics*, vol. 76, p. 102169, ISSN 0927-5371, doi:10.1016/j.labeco.2022.102169.

Barrero, J.M., Bloom, N. and Davis, S.J. (2021). 'Why Working From Home Will Stick', NBER Working Paper 28731, National Bureau of Economic Research, doi:10.3386/w28731.

Barrero, J.M., Bloom, N. and Davis, S.J. (2023). 'The Evolution of Work from Home', *Journal of Economic Perspectives*, vol. 37(4), pp. 23–49, ISSN 0895-3309, doi:10.1257/jep.37.4.23.

Bloom, N., Han, R. and Liang, J. (2024). 'Hybrid working from home improves retention without damaging performance', *Nature*, pp. 1–6, ISSN 1476-4687, doi:10.1038/s41586-024-07500-2.

Bloom, N., Liang, J., Roberts, J. and Ying, Z.J. (2015). 'Does Working from Home Work? Evidence from a Chinese Experiment', *The Quarterly Journal of Economics*, vol. 130(1), pp. 165–218, ISSN 0033-5533, 1531-4650, doi:10.1093/qje/qju032.

Coskun, S., Dauth, W., Gartner, H., Stops, M. and Weber, E. (2024). 'Working from Home Increases Work-Home Distances', IZA Discussion Paper 16855, Institute of Labor Economics.

Delventhal, M. and Parkhomenko, A. (2023). 'Spatial Implications of Telecommuting', SSRN Working Paper, Social Science Research Network, doi:10.2139/ssrn.3746555.

Emanuel, N., Harrington, E. and Pallais, A. (2023). 'The Power of Proximity to Coworkers: Training for Tomorrow or Productivity Today?', doi:10.3386/w31880.

Goldin, C. (2014). 'A Grand Gender Convergence: Its Last Chapter', *American Economic Review*, vol. 104(4), pp. 1091–1119, ISSN 0002-8282, doi:10.1257/aer.104.4.1091.

Goldin, C. (2022). 'Understanding the Economic Impact of COVID-19 on Women', *Brookings Papers on Economic Activity*, pp. 65–110.

Gupta, A., Mittal, V., Peeters, J. and Van Nieuwerburgh, S. (2022). 'Flattening the Curve: Pandemic-Induced Revaluation of Urban Real Estate', *Journal of Financial Economics*, vol. 146(2), pp. 594–636, doi:10.1016/j.jfineco.2021.10.008.

Hall, A., Hünefeld, L. and Rohrbach-Schmidt, D. (2020). 'BIBB/BAuA-Erwerbstätigenbefragung 2018 - Arbeit und Beruf im Wandel', doi:10.7803/501.18.1.1.10.

Hangartner, D., Marbach, M., Henckel, L., Maathuis, M.H., Kelz, R.R. and Keele, L. (2021). 'Profiling Compliers in Instrumental Variables Designs', doi:10.48550/arXiv.2103.06328.

Harrington, E. and Kahn, M.E. (2023). 'Has the Rise of Work-from-Home Reduced the Motherhood Penalty in the Labor Market?', Tech. rep.

Imbens, G.W. and Rubin, D.B. (2015). *Causal Inference for Statistics, Social, and Biomedical Sciences: An Introduction*, Cambridge: Cambridge University Press, ISBN 978-0-521-88588-1, doi:10.1017/CBO9781139025751.

Jayachandran, S., Nassal, L., Notowidigdo, M.J., Paul, M., Sarsons, H. and Sundberg, E. (2024). 'Moving to Opportunity, Together', NBER Working Paper 32970, National Bureau of Economic Research, doi:10.3386/w32970.

Lewis, D.J., Mertens, K. and Federal Reserve Bank of Dallas (2024). 'A Robust Test for Weak Instruments for 2SLS with Multiple Endogenous Regressors', Federal Reserve Bank of Dallas Working Paper 2208, FED of Dallas.

Lyttelton, T., Zang, E. and Musick, K. (2020). ‘Gender Differences in Telecommuting and Implications for Inequality at Home and Work’, SSRN Scholarly Paper 3645561, Social Science Research Network, Rochester, NY.

Marbach, M. and Hangartner, D. (2020). ‘Profiling Compliers and Noncompliers for Instrumental-Variable Analysis’, *Political Analysis*, vol. 28(3), pp. 435–444, ISSN 1047-1987, 1476-4989, doi:10.1017/pan.2019.48.

Mas, A. and Pallais, A. (2017). ‘Valuing Alternative Work Arrangements’, *American Economic Review*, vol. 107(12), pp. 3722–3759, ISSN 0002-8282, doi:10.1257/aer.20161500.

Nagler, M., Rincke, J. and Winkler, E. (2024). ‘Working from home, commuting, and gender’, *Journal of Population Economics*, vol. 37(3), p. 58, ISSN 1432-1475, doi:10.1007/s00148-024-01035-6.

Nennstiel, R. and Hudde, A. (2025). ‘Is there a growing gender divide among young adults in regard to ideological left–right self-placement? Evidence from 32 European countries’, *European Sociological Review*, p. jcaf021, ISSN 0266-7215, doi:10.1093/esr/jcaf021.

Olea, J.L.M. and Pflueger, C. (2013). ‘A Robust Test for Weak Instruments’, *Journal of Business & Economic Statistics*, vol. 31(3), pp. 358–369, ISSN 0735-0015.

Oswald, A.J. and Powdthavee, N. (2010). ‘Daughters and Left-Wing Voting’, *The Review of Economics and Statistics*, vol. 92(2), pp. 213–227, ISSN 0034-6535, doi:10.1162/rest.2010.11436.

Pabilonia, S.W. and Vernon, V. (2023). ‘Who is doing the chores and childcare in dual-earner couples during the COVID-19 era of working from home?’, *Review of Economics of the Household*, vol. 21(2), pp. 519–565, ISSN 1573-7152, doi:10.1007/s11150-022-09642-6.

Schüller, S. (2025). ‘Estimating the Effect of Working from Home on Parents’ Division of Childcare and Housework: A New Panel IV Approach’, CESifo Working Paper 11689.

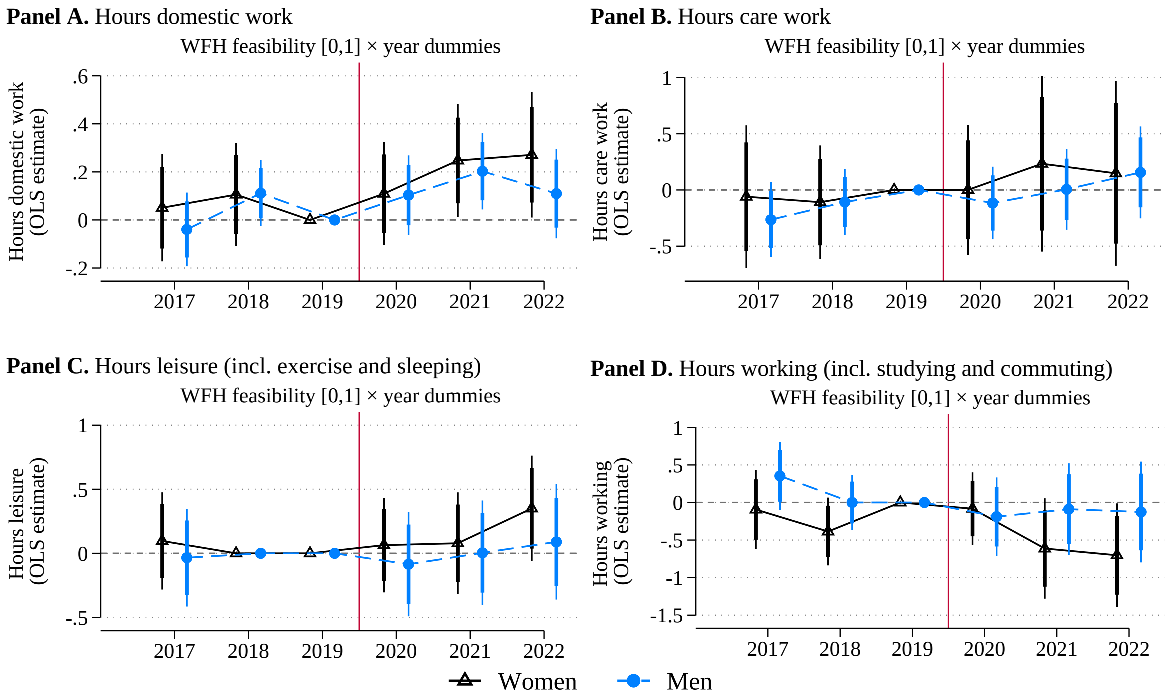
von Gaudecker, H.M., Holler, R., Simon, L. and Zimpelmann, C. (2024). ‘Can Work from Home Help Balance the Parental Division of Labor?’, ECONtribute Discussion Paper 321.

Appendix

Does Remote Work Reinforce Gender Gaps in (Un)Paid Labor?
by *Jean-Victor Alipour*

Appendix A Figures

Figure A.1: ITT effects of WFH on individual time use on an average weekday, balance 2017–2022



Notes: The figure plots DiD estimates based on Equation 1. The sample includes individuals employed and aged 16–64 in 2019, balanced between 2017–2022. 99% and 95%-confidence intervals are drawn using standard errors clustered by individual. Data are from the SOEP.

Appendix B Tables

Table B.1: Summary statistics (couple level)

	(1) Women N = 875	(2) Men N = 875	(3) Within-couple gender gap (W-M)
A. WFH variables			
WFH at least 1x/week, 2022	0.283 (0.451)	0.323 (0.468)	-0.040 (0.568)
Weekly hours WFH, 2022 (all)	5.329 (10.588)	8.411 (14.220)	-3.083 (16.197)
Weekly hours WFH, 2022 (only WFH 1+/week)	18.373 (12.217)	24.802 (14.159)	-8.137 (16.766)
Share of hours WFH, 2022 (all)	0.181 (0.327)	0.200 (0.338)	-0.019 (0.421)
Share of hours WFH, 2022 (only WFH 1+/week)	0.561 (0.347)	0.591 (0.333)	-0.064 (0.446)
WFH feasibility (0/1), 2019	0.506 (0.500)	0.509 (0.500)	-0.002 (0.592)
WFH feasibility [0,1], 2019	0.372 (0.290)	0.370 (0.296)	0.002 (0.339)
B. Main outcome variables			
Hours leisure (incl. exercise and sleeping), 2019	8.498 (1.632)	8.295 (1.647)	0.203 (1.913)
Δ Hours leisure (incl. exercise and sleeping)	0.124 (1.769)	0.329 (2.116)	-0.205 (2.411)
Hours working (incl. studying and commuting), 2019	7.019 (2.625)	9.352 (1.771)	-2.333 (3.158)
Δ Hours working (incl. studying and commuting)	-0.549 (3.183)	-1.094 (3.075)	0.545 (4.132)
Hours care work, 2019	2.571 (3.627)	1.057 (1.654)	1.514 (3.061)
Δ Hours care work	0.144 (4.482)	0.089 (1.934)	0.055 (3.952)
Hours domestic work (incl. errands), 2019	2.825 (1.321)	1.437 (1.028)	1.389 (1.727)
Δ Hours domestic work (incl. errands)	0.271 (1.508)	0.502 (1.562)	-0.231 (2.065)
Hours DIY activities, 2019	0.401 (0.582)	0.582 (0.805)	-0.181 (0.877)
Δ Hours DIY activities	0.321 (0.769)	0.472 (1.019)	-0.151 (1.183)
Weekly hours worked, 2019	30.316 (11.801)	41.822 (8.194)	-11.505 (14.677)
Δ Weekly hours worked	-2.737 (13.289)	-5.151 (14.340)	2.414 (17.784)

Notes: The table reports variable means and standard deviations (in parentheses) for female and male partners of the same couple. Column 3 reports within-couple gender differences. The sample includes couples in which both partners are employed and aged 16–64 in 2019, and who completed the time-use questionnaire in 2019 and 2022. The operator Δ denotes changes between 2019 and 2022. The data are from the SOEP.

Table B.2: 2SLS estimates of the impact of WFH (hours per weekday) on individual time use

	(1)	(2)	(3)	(4)	(5)	(6)
	Δ Hours spent on average weekday on					
	Dom. work (incl. errands)	DIY activities	Care work	Leisure (incl. sleep, exercise)	Work (incl. trainings, commuting)	Paid work
Panel A: Women						
Hours WFH per weekday, 2022	0.104***	0.018	0.069	0.124***	-0.273***	-0.099*
	(0.032)	(0.017)	(0.080)	(0.040)	(0.069)	(0.059)
#Individuals	2,992	2,992	2,992	2,992	2,992	2,992
2019 DV mean	2.59	.373	1.923	8.541	7.47	6.405
Prop. effect at the mean	.04	.047	.036	.015	-.037	-.016
First-stage coef.	3.13	3.13	3.13	3.13	3.13	3.13
	[543.86]	[543.86]	[543.86]	[543.86]	[543.86]	[543.86]
Panel B: Men						
Hours WFH per weekday, 2022	0.039**	0.018	0.023	0.020	-0.020	0.005
	(0.020)	(0.015)	(0.026)	(0.030)	(0.044)	(0.041)
#Individuals	2,513	2,513	2,513	2,513	2,513	2,513
2019 DV mean	1.496	.473	.814	8.466	9.308	8.244
Prop. effect at the mean	.026	.038	.029	.002	-.002	.001
First-stage coef.	4.78	4.78	4.78	4.78	4.78	4.78
	[803.21]	[803.21]	[803.21]	[803.21]	[803.21]	[803.21]
p-values test of equal effects	0.0795	0.9809	0.5846	0.0364	0.0019	0.1429

Notes: The table presents 2SLS results based on Equations 2 and 3. The explanatory variable of interest is the average number of hours worked from home per weekday in 2022 (weekly hours WFH divided by five). The instrument is the WFH feasibility [0,1] of individuals' 2019 job (see Section 2.2 for details). The sample includes individuals employed and aged 16–64 in 2019, who are observed again in 2022. Kleibergen-Paap F-statistic for the first-stage regression in brackets. Standard errors are clustered by individual and reported in parentheses. Data are from the SOEP. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$

Table B.3: 2SLS estimates of the impact of WFH (at least once per week) on individual time use

	(1)	(2)	(3)	(4)	(5)	(6)
	Δ Hours spent on average weekday on					
	Dom. work (incl. errands)	DIY activities	Care work	Leisure (incl. sleep, exercise)	Work (incl. trainings, commuting)	Paid work
Panel A: Women						
WFH at least 1x/week in 2022	0.457***	0.071	0.167	0.620***	-1.251***	-0.538**
	(0.146)	(0.079)	(0.372)	(0.189)	(0.329)	(0.274)
#Individuals	2,992	2,992	2,992	2,992	2,992	2,992
2019 DV mean	2.59	.373	1.923	8.541	7.47	6.405
Prop. effect at the mean	.177	.189	.087	.073	-.167	-.084
First-stage coef.	0.39	0.39	0.39	0.39	0.39	0.39
	[738.55]	[738.55]	[738.55]	[738.55]	[738.55]	[738.55]
Panel B: Men						
WFH at least 1x/week in 2022	0.193	0.145*	0.151	0.111	-0.045	0.037
	(0.118)	(0.087)	(0.154)	(0.174)	(0.258)	(0.236)
#Individuals	2,513	2,513	2,513	2,513	2,513	2,513
2019 DV mean	1.496	.473	.814	8.47	9.308	8.244
Prop. effect at the mean	.129	.306	.186	.013	-.005	.004
First-stage coef.	0.48	0.48	0.48	0.48	0.48	0.48
	[926.31]	[926.31]	[926.31]	[926.31]	[926.31]	[926.31]
p-values test of equal effects	0.1597	0.5295	0.9694	0.0477	0.0040	0.1124

Notes: The table presents 2SLS results based on Equations 2 and 3, estimated separately by gender. The explanatory variable of interest is a dummy identifying individuals who WFH at least once per week in 2022. The instrument is the WFH feasibility (0/1) of individuals' 2019 job (see Section 2.2 for details). The sample includes individuals employed and aged 16–64 in 2019, who are observed again in 2022. Kleibergen-Paap F-statistic for the first-stage regression in brackets. Standard errors are clustered by individual and reported in parentheses. Data are from the SOEP. $^{***}p < 0.01,^{**}p < 0.05,^{*}p < 0.1$

Table B.4: 2SLS estimates of the impact of WFH (hours per weekday) on work and living conditions

	(1) Moved (0/1)	(2) #Rooms	(3) Log living space	(4) Left employm. (0/1)	(5) Employer change (0/1)	(6) Log gross labor income	(7) Log gross hrly wage	(8) Married (0/1)	(9) Single (0/1)	(10) Children <14 (0/1)
Panel A: Women										
Hours WFH per weekday in 2022	-0.001	0.069***	0.009*	-0.020***	-0.028***	-0.039***	-0.015	0.001	-0.006*	0.006
	(0.006)	(0.025)	(0.005)	(0.006)	(0.008)	(0.011)	(0.010)	(0.005)	(0.004)	(0.008)
#Individuals	2,992	2,972	2,992	2,992	2,624	2,624	2,607	2,991	2,991	2,992
DV mean	.103	4.376	4.634	.123	.186	7.575	2.748	.586	.252	.355
First-stage coef.	3.13	3.14	3.13	3.13	3.47	3.43	3.45	3.13	3.13	3.13
	[543.9]	[544.0]	[543.9]	[543.9]	[575.2]	[564.1]	[569.2]	[544.0]	[544.0]	[543.9]
Panel B: Men										
Hours WFH per weekday in 2022	0.012**	0.039*	0.008**	-0.008*	-0.006	0.003	0.012*	0.001	-0.002	0.002
	(0.005)	(0.020)	(0.004)	(0.004)	(0.005)	(0.006)	(0.007)	(0.004)	(0.003)	(0.005)
#Individuals	2,513	2,488	2,513	2,513	2,270	2,269	2,256	2,511	2,511	2,513
DV mean	.111	4.38	4.632	.097	.145	8.122	2.979	.641	.275	.353
First-stage coef.	4.78	4.78	4.78	4.78	5.21	5.18	5.23	4.78	4.78	4.78
	[803.2]	[796.3]	[803.2]	[803.2]	[878.8]	[863.5]	[875.3]	[803.0]	[803.0]	[803.2]
p-values test of equal effects	0.1055	0.3404	0.7696	0.1100	0.0132	0.0004	0.0205	0.9725	0.3593	0.5955

Notes: The table presents 2SLS results based on Equations 2 and 3, estimated separately by gender. The instrument is the WFH feasibility [0,1] of individuals' 2019 job (see Section 2.2 for details). The sample includes individuals employed and aged 16–64 in 2019, who are observed again in 2022. Columns 5–7 exclude individuals who are not employed in 2022. Kleibergen-Paap F-statistic for the first-stage regression in brackets. Standard errors are clustered by individual and reported in parentheses. Data are from SOEP. $^{***}p < 0.01, ^{**}p < 0.05, ^{*}p < 0.1$

Table B.5: 2SLS estimates of the impact of WFH (hours per weekday) on couples' gender time-use gaps

	(1)	(2)	(3)	(4)	(5)	(6)
			Δ Time use for			
	Dom. work (incl. errands)	DIY activities	Care work	Leisure (incl. sleep, exercise)	Work (incl. trainings, commuting)	Paid work
Panel A: Effect on female partner's time use						
Female partner's hours WFH, 2022	0.151**	-0.042	0.316	-0.012	-0.285**	-0.123
	(0.069)	(0.036)	(0.207)	(0.076)	(0.142)	(0.121)
Male partner's hours WFH, 2022	-0.030	0.021	0.151	-0.011	-0.016	-0.030
	(0.038)	(0.021)	(0.123)	(0.047)	(0.085)	(0.071)
#Couples	875	875	875	875	875	875
Combined effect ($\hat{\gamma}_1 + \hat{\gamma}_2$)	.121*	-.021	.467**	-.024	-.301**	-.153
	(.062)	(.031)	(.183)	(.069)	(.131)	(.112)
2019 DV mean	2.825	.401	2.571	8.498	7.019	6.063
Panel B: Effect on male partner's time use						
Female partner's hours WFH, 2022	-0.038	-0.058	-0.007	-0.014	0.023	0.070
	(0.062)	(0.040)	(0.087)	(0.092)	(0.125)	(0.114)
Male partner's hours WFH, 2022	0.077**	0.041*	0.137***	-0.017	0.015	0.101
	(0.034)	(0.025)	(0.051)	(0.055)	(0.075)	(0.070)
Combined effect ($\hat{\gamma}_1 + \hat{\gamma}_2$)	.039	-.017	.13*	-.031	.039	.171
	(.069)	(.041)	(.078)	(.091)	(.121)	(.113)
2019 DV mean	1.437	.582	1.057	8.295	9.352	8.364
Panel C: Effect on couple's time-use gap (W-M)						
Female partner's hours WFH, 2022	0.189**	0.016	0.322*	0.002	-0.309*	-0.194
	(0.089)	(0.051)	(0.180)	(0.106)	(0.175)	(0.150)
Male partner's hours WFH, 2022	-0.107**	-0.021	0.014	0.006	-0.031	-0.131
	(0.050)	(0.031)	(0.108)	(0.064)	(0.107)	(0.089)
Combined effect ($\hat{\gamma}_1 + \hat{\gamma}_2$)	.082	-.004	.337**	.007	-.34**	-.325**
	(.089)	(.046)	(.157)	(.1)	(.165)	(.145)
Average effect ($\hat{\gamma}_1 \times \overline{WFH22}^f + \hat{\gamma}_2 \times \overline{WFH22}^m$)	.022	-.017	.368**	.011	-.382*	-.427**
	(.108)	(.056)	(.187)	(.122)	(.201)	(.177)
2019 DV mean	1.389	-.181	1.514	.203	-2.333	-2.301
#Couples	875	875	875	875	875	875
First-stage diagnostics						
Lewis-Mertens statistic (g_{min})	67.60	67.60	67.60	67.60	67.60	67.60
g_{min} critical values ($\alpha = 0.05, \tau = 0.1$)	20.63	20.63	20.63	20.63	20.63	20.63

Notes: The table presents 2SLS results based on Equations 4–6. Time-use gaps (Panel C) refer to the female-male differences in hours spent on a given activity on an average weekday within a couple. The instruments are the WFH feasibility [0,1] of each partner's 2019 job (see Section 2.2 for details). The sample includes couples in which both partners are employed and aged 16–64 in 2019, and observed again in 2022. Standard errors are clustered by couple and reported in parentheses. Data are from the SOEP. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$

Table B.6: Complier analysis

	(1) Sample	(2) Complier	(3) Never taker	(4) Always taker	(5) Difference (Sample-Complier)	(6) Std. difference (Sample-Complier)
Panel A: Women						
Age, 2019	45.42	44.02	47.05	39.71	1.40***	0.14
	(0.19)	(0.52)	(0.33)	(1.20)	(0.49)	
Partner in HH (0/1), 2019	0.69	0.65	0.71	0.70	0.04*	0.09
	(0.01)	(0.02)	(0.02)	(0.05)	(0.02)	
Children under 14 in HH (0/1), 2019	0.35	0.38	0.32	0.47	-0.03	-0.06
	(0.01)	(0.02)	(0.01)	(0.06)	(0.02)	
No of children u14 in HH, 2019	0.54	0.57	0.49	0.84	-0.03	-0.03
	(0.02)	(0.04)	(0.03)	(0.10)	(0.04)	
Commuting distance (km), 2019	14.10	12.95	14.23	19.95	1.14	0.06
	(0.34)	(0.93)	(0.63)	(3.61)	(0.88)	
Weekly hours worked, 2019	32.03	30.60	33.29	29.65	1.42***	0.11
	(0.23)	(0.59)	(0.41)	(1.47)	(0.53)	
Gross monthly wage (EUR), 2019	2489.63	2179.25	2717.29	2398.71	310.38***	0.21
	(27.31)	(76.39)	(59.35)	(156.70)	(73.67)	
Firm tenure in years, 2019	10.63	7.91	12.68	9.34	2.72***	0.29
	(0.17)	(0.55)	(0.38)	(1.12)	(0.53)	
Firm with 100+ employees (0/1), 2019	0.62	0.63	0.62	0.63	-0.01	-0.02
	(0.01)	(0.02)	(0.02)	(0.05)	(0.02)	
Panel B: Men						
Age, 2019	45.40	43.44	47.96	43.07	1.95***	0.19
	(0.21)	(0.49)	(0.45)	(1.08)	(0.43)	
Partner in HH (0/1), 2019	0.76	0.76	0.76	0.70	-0.00	-0.00
	(0.01)	(0.02)	(0.02)	(0.05)	(0.01)	
Children under 14 in HH (0/1), 2019	0.35	0.38	0.32	0.36	-0.03*	-0.07
	(0.01)	(0.02)	(0.02)	(0.06)	(0.02)	
No of children u14 in HH, 2019	0.59	0.63	0.53	0.63	-0.04	-0.05
	(0.02)	(0.04)	(0.04)	(0.11)	(0.03)	
Commuting distance (km), 2019	18.22	20.47	15.64	19.69	-2.25***	-0.12
	(0.40)	(0.77)	(0.61)	(2.05)	(0.59)	
Weekly hours worked, 2019	41.22	40.87	41.81	40.02	0.35	0.04
	(0.19)	(0.38)	(0.38)	(1.10)	(0.35)	
Gross monthly wage (EUR), 2019	4061.22	3666.81	4487.92	4104.53	394.41***	0.17
	(46.02)	(109.95)	(117.64)	(219.84)	(102.91)	
Firm tenure in years, 2019	11.97	10.28	13.92	11.49	1.69***	0.17
	(0.20)	(0.42)	(0.50)	(1.08)	(0.41)	
Firm with 100+ employees (0/1), 2019	0.69	0.69	0.67	0.74	-0.01	-0.01
	(0.01)	(0.02)	(0.02)	(0.04)	(0.02)	

Notes: The table reports covariate means for the whole sample and subsamples of compliers, never-takers, and always takers based on the methodology by Marbach and Hangartner (2020); Hangartner *et al.* (2021). Column 5 shows differences in means between the sample and compliers. Column 6 shows standardized differences. The instrument or "treatment assignment" is WFH feasibility (0/1), and the treatment is WFH at least once per week in 2022 (0/1). Compliers work remotely if and only if assigned to treatment. Never-takers (always-takers) refuse (obtain) treatment irrespective of assignment. The sample includes individuals employed and aged 16–64 in 2019, who are observed again in 2022. Among women, 39% of individuals are identified as compliers, 55% are never-takers, and 6% are always-takers. Among men, the proportions are 48% compliers, 44% never-takers, and 8% always-takers. Bootstrapped standard errors (1,000 replications) in parentheses. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$