

Matching it up: working arrangements and job satisfaction^{*}

Katarzyna Bech-Wysocka

FAME|GRAPE and Warsaw School of Economics

Magdalena Smyk

FAME|GRAPE and Warsaw School of Economics

Joanna Tyrowicz

University of Regensburg, FAME| GRAPE, and IZA[†]

Lucas van der Velde

FAME|GRAPE and Warsaw School of Economics

September 2021

Abstract

We leverage the flexibility enactment theory to study empirically the link between working arrangements and job satisfaction. We provide novel insights on the match between the individual inclination to work in non-standard working arrangements and the factual conditions of employment. We thus reconcile the earlier literature, which found both positive and negative effects of non-standard employment on job satisfaction. Using data from the European Working Conditions Survey we characterize the extent of mismatch between individual inclination and factual working arrangements. We provide several novel results. First, the extent of mismatch is substantial and reallocating workers between jobs could substantially boost overall job satisfaction in European countries. Second, the mismatch more frequently plagues women and parents. Finally, we demonstrate that the extent of mismatch is heterogeneous across countries, which shows that one-size-fits-all policies are not likely to maximize the happiness of workers, whether flexibility is increased or reduced.

Key words: working time flexibility, job satisfaction, gender

^{*}The authors express gratitude for insightful comments and suggestions by Julia McQuillan, Jekaterina Navicke, Erita Narhetali, and Linas Tarasonis. Kazimierz Czarnocki and Kamil Sekut provided wonderful research assistance. Financial support by National Science Center grant no. #2017/27/L/HS4/O3219 is much appreciated. The remaining errors are ours.

[†]Corresponding author: j.tyrowicz@grape.org.pl

1 Introduction

The prevailing narrative in public debate is that “flexible worker is a happy worker”.¹ This narrative draws on multiple empirical studies, where employers agreed to alter working time arrangements in response to preferences of the workers. Such arrangements allow for greater work-life balance and are believed to contribute not only to workers’ happiness (Atkinson et al., 2011) but also to their productivity (Bloom et al., 2015).

Against this narrative, sociological theory of the working time flexibility provides a highly complex portrait. Specifically, the *flexibility enactment theory* proposes to view the flexibility through the lenses of boundary management strategy (Kossek et al., 2004). Building on Lee et al. (2002), this theory posits that not everybody is equally suited to work in a flexible environment. This conceptualization implies that the same individual may prefer some forms of flexibility over others. It also implies that individual conditions are reflected in the impact of a given form of flexibility of life and job satisfaction.

Indeed, flexible working hours and ability to work from outside the official premises creates room for work-related matters to invade the private spheres, both in terms of time and in terms of behavior.² The digitization of work-related communication, companies operations spanning multiple time zones and other technological changes imply that many workers report anxiety about being expected to be reachable 24/7, and about work transgressing their private spaces, triggering policy initiatives such as *the right to disconnect*.³

The empirical evidence to date suggests that access to flexible working time arrangements itself does not directly translate to life and job satisfaction. Generally, individuals with caring obligations appreciate if the employer somehow accommodates for their complicated time schedules (Bainbridge and Townsend, 2020). Indeed, full-time work makes mothers less happy than part-time or staying home altogether in Europe (Hamplová, 2019). However, the spillovers are highly heterogeneous across the types of flexibility and gender, with multitude of patterns emerging from the data (see Lott, 2015, 2020, for a cross-country comparative study, and for uniquely profound data and analysis for Germany, respectively). The results are indeed mixed. While workplaces that discourage flexibility are considered less attractive by all types of workers (O’Connor and Cech, 2018), the same does not hold for the evaluation of workers actually making use of these instruments. Workers of both genders, who ask for flexible work arrangements, are evaluated negatively (Cech and Blair-Loy, 2014; Munsch, 2016) and as poor workers (Rudman and Mescher, 2013; Vandello et al., 2013).

Ambiguity haunts also the empirical findings concerning job satisfaction and working in flexible arrangements. Wheatley (2017) finds positive correlations for men, but negative for women. Bell-

¹One typical example is this [article from Forbes](#).

²For example, an employee may enjoy waiting at home when on call rather than be on duty at the premises. However, if the call can occur at irregular hours, including night hours, working from home interferes with the lives of other family members, effectively deteriorating life satisfaction.

³In Europe, the *Working Time Directive* specifies that the time when the worker is expected to answer the phone calls or emails should count as a regular working time, irrespective of whether the phones/emails occurred. France and Italy have introduced specific legislation, the former requires that all employment contracts specify which hours are “off grid” for the employers, in the latter the legislation concerns self-employed ICT-based workers. The US perspective is covered extensively by Secunda (2019).

mann and Hübler (2020) argue that the patterns for job satisfaction are generally unclear, while the correlations with work-life balance are robustly negative. Hayman (2009) reports lower level of work-family and family-work conflicts, especially for women, but extant literature argues that the key is the attitude towards flexibility among HR workers and direct managers (e.g. McCarthy et al., 2013), which, as discussed above, is typically negative. Moreover, flexible hours may imply cognitive spillovers to leisure, deteriorating health and increasing the psychological strain of work (e.g. Williams et al., 2013; Cha and Weeden, 2014).

Building on the flexibility enactment theory, we theorize that at a given point in time, each of us has an individual *inclination* to work in non-standard employment forms (NSE). For example, a young couple without children and caring obligations may appreciate long working hours if they view this arrangement as a fast-track to career, but those same individuals may no longer find this form of NSE rewarding when they become parents. The key hypotheses of our study are the following.

Hypothesis 1 *Countries are heterogeneous in the degree of mismatch between preferred and actual NSE forms; in all countries a reallocation of workers in a way that reduces these mismatches would raise aggregate job satisfaction.*

Hypothesis 2 *Mismatches between actual and preferred NSE forms in working arrangements exist and are systematic: they affect women and parents to a larger extent than other groups of workers.*

Admittedly, inclination to work in flexible arrangements is unobservable. We propose to employ machine learning methods to uncover this latent link between job satisfaction and NSE. We study different types of flexibility. Once we reveal the relationship between job satisfaction and (the types of) work flexibility, we utilize these insights to identify the individuals who are mismatched and the individuals who are matched. Specifically, we recover the gradient of job satisfaction to (the types of) work flexibility and thus we can identify the workers who *would benefit* from non-standard working arrangements, but do not have them and, conversely, the workers who would benefit from standard working arrangements, but reported non-standard ones. Finally, we characterize this mismatch across individual characteristics and across countries.

This paper aims to make several contributions to the existing literature. First, to the best of our knowledge, we are the first to operationalize the flexibility enactment theory in quantitative empirical context. We do so by introducing machine learning methods *in lieu* of standard parametric approach, which is our second contribution. This choice allows exploiting fully the informational content of the data without imposing any ad hoc restrictions in terms of model specification and functional forms. Thus, while our approach is driven by theory, we avoid arbitrary choices in terms of modeling and functional form. The advantage of this method is also that we obtain directly the model-implied counterfactual job satisfaction levels rather than parametrically estimated predictions on which we would need to impose arbitrary boundaries to recover the levels of job satisfaction. Third, our analysis is set in an international, comparative context. By using the data from the European Working Conditions Survey, we can explore several welfare regimes and search the links from legal and cultural contexts to the mismatch between the actual and the optimal working arrangements.

Our study is structured as follows. Section 2 provides an overview of the existing literature. In section 3 we present the data and in section 4 we introduce our empirical approach. The results of our estimations as well as the counterfactual experiments are discussed at length in section 5. Exploring heterogeneity of the mismatch between actual and optimal working arrangements, we discuss in detail the drivers of this mismatch as well as country-specific implications.

2 Literature review

Sociological theory, early on, provided several explanations for why individual inclination to non-standard working arrangements may be highly heterogeneous. First, the *role theory* and the *self-discrepancy theory* and the *social identity theory* all emphasize that individuals define themselves through the lenses of how they are perceived by others and by themselves. These lenses are related to individual personality characteristics and interpersonal relations (Kahn et al., 1964; Katz and Kahn, 1978), they are by design individual and thus highly heterogeneous in a sense that the social norms and stereotypes exhibit differently across social classes, education achievement levels or even local communities. This combination of individual own perceptions and perceptions of general social norms naturally alleviates or exacerbates the work-family and family-work conflicts. This line of scholarly work was further advanced by the *spillover theory* (Westman and Piotrkowski, 1999) and the *boundary theories* (Nippert-Eng, 1996; Clark, 2000). With their specific angles, these theories posit that integrating life experiences from both the personal and the work environments, people automatically transmit these experiences across the life spheres and thus spillovers are unavoidable. These spillovers are not limited to negative consequences: one sphere can and often does enrich the other.

However, the spillovers ought to be actively managed. The crux of the *flexibility enactment theory* is that the ability to actively manage the spillovers and thus the ability to derive satisfaction from a job of given working arrangements depends on an individual ability to manage boundaries overall (Kossek et al., 2004). The key individual characteristic is that some individuals find it easier to manage the boundaries between work and life, whereas for the others this demarcation is more troublesome. An individual boundary management strategy relates to a combination of boundaries (related to time, space, sense of belonging, etc.) and it is not defined as dichotomous over two states (flexible or not), rather it is a continuum.⁴

Despite this rich body of theoretical work in psychology, social psychology and sociology, the empirical treatment of non-standard working arrangements focuses typically on one aspect of working time arrangements and one specific characteristic of the job contract. For example, it was repeatedly established that full-time work makes mothers less happy than part-time or staying home altogether (see Hamplová, 2019, for evidence from Europe). This result extends to all individuals with caring obligations, who appreciate if the employer somehow accommodates for

⁴In the remainder of this article we consider all forms of irregular or non-standard employment as “flexible”, that is this term encompasses part-time, long hours, weekends, nights, ability to set own hours (or, rather, lack thereof). We also account for negative work amenities which may entangle the management of boundaries, such as long commute, the need to be available on call, etc.

complications in terms of time availability implied by those obligations (Bainbridge and Townsend, 2020, for evidence from Australia).

The qualitative and quantitative evidence supports the hypothesis that job pressure is felt strongly in occupations where time boundaries are harder to set and place boundaries are not enough to establish them. A study by Schieman and Glavin (2016) demonstrates empirical evidence to corroborate this observation exploring a nationally representative sample of workers and delineating the job pressure for individuals holding high status jobs, such as professionals and managers. In a study comparing workers from four culturally distant countries, Barney and Elias (2010) argues that in some cultural contexts, the autonomy in setting one's own schedule actually exacerbates stress related to work and affect intrinsic motivation. Dumas and Sanchez-Burks (2015) argue in an extensive review paper that the perspective on boundary management and spillovers in the literature alternates between treating them as a tool for handling role responsibilities on the one hand and as a tool for shaping workplace identity and relationships on the other.

Finally, growing body of experimental evidence shows that workers' appraisal of flexible working time arrangements is heterogeneous across groups and types of flexibility (Mas and Pallais, 2017). There do not seem to be differences in preference for long hours, rather differences in time endowments, i.e. ability to work long hours (Cortes and Pan, 2019; Zapf and Weber, 2017) or specific hours (Duchini and van Effenterre, 2018; Cubas et al., 2019).

There is also ample qualitative evidence that men less frequently request flexibility (Vandello et al., 2013), whereas requesting flexibility is stigmatizing (Rudman and Mescher, 2013; Vandello et al., 2013; Munsch, 2016) and stereotypically associated with different motivations across genders: leisure aspirations for men and family devotion schema for women (Williams et al., 2013).⁵ Indeed, men are less likely to have their flexibility request accepted if it is for family reasons (Brescoll et al., 2013). Requesting flexible working time arrangements opposes the "work devotion schema" (Williams et al., 2013), because the devoted and committed worker is invariably available to the employer and does not need the employer to accommodate for personal complications in terms of time schedules (Collewet et al., 2017). Both the work-devotion schema and the family-devotion schema are implicitly opposing the flexibility enactment theory in that they allow only extremes rather than coexistence of both family and work life, with managed boundaries.

Against this rich and informative literature, we propose several innovations. First, rather than hypothesizing about the type of workers and the types of occupations that may be more challenged by boundary management, we focus on identifying the *gradient* of job satisfaction across all types of non-standard working arrangements. This gradient reflects a hypothesis long salient in the literature, that a given person may have derived different levels of job satisfaction, if some specific features of this job were altered. For some individuals, this gradient may be increasing in irregular working arrangements, which is to say that those individuals are better off with irregular rather than with standard working arrangements. The opposite may hold true for some other individuals. In short, working time arrangements have an effect on job satisfaction through the mismatch to

⁵This creates a self-fulfillment prophecy, where firm expectations about household division of labor leads to realization of such division, even if ex-ante there was no particular advantage over alternative arrangements (see Albanesi and Olivetti, 2009; Flabbi and Moro, 2012).

the workers' needs rather than by the arrangements themselves. This idea of gradient builds on the earlier conceptualization of the self-discrepancy theory and social identity theory, i.e. for some individuals irregularity is a natural habitat, whereas for the others it triggers negative spillovers and transgressions, necessitating burdensome development of a boundary management strategy.

Second, we make no *ex ante* assumptions about the drivers of the gradient. We employ machine learning rather than parametric modelling. This methodology allows us to avoid many discretionary choices. Foremost, we do not need to impose any restrictions on the functional form of the model nor on the statistical distribution of the data. Neither the definition of variables nor their interrelationships have to be pre-specified in model which relies solely on classification. Moreover, given that job satisfaction is typically a categorical variable, we do not need to assume anything about the thresholds for assigning the parametric predictions to the categories: the machine learning classifications automatically assign the level of job satisfaction to each individual.

Third, we explore the data from the representative cross-country European Working Conditions Survey. Naturally, this choice makes it impossible for us to focus on some special groups such as high-aspiration professionals and managers. However, this data permits the identification of the gradient and the mismatch along all occupational groups and industries. With this choice, our analysis has lower risk of omitting some important pockets of mismatch between the actual and the ideal working arrangements, yielding fruitful grounds for further research into the topic.

3 Data

We rely on data from the European Working Conditions Survey spanning years between 1991 and 2015. EWCS is administered across countries, with harmonized sampling methodology as well as questionnaire. This survey is administered every five years. The sample currently covers 36 countries⁶ and nearly 180 thousand individuals. With few exceptions, samples within countries comprise roughly 1000 individuals (for Germany across all waves and in selected waves for the other countries, the samples are several times larger). Given this disparity, all our specifications account for sample weights and country-fixed effects.

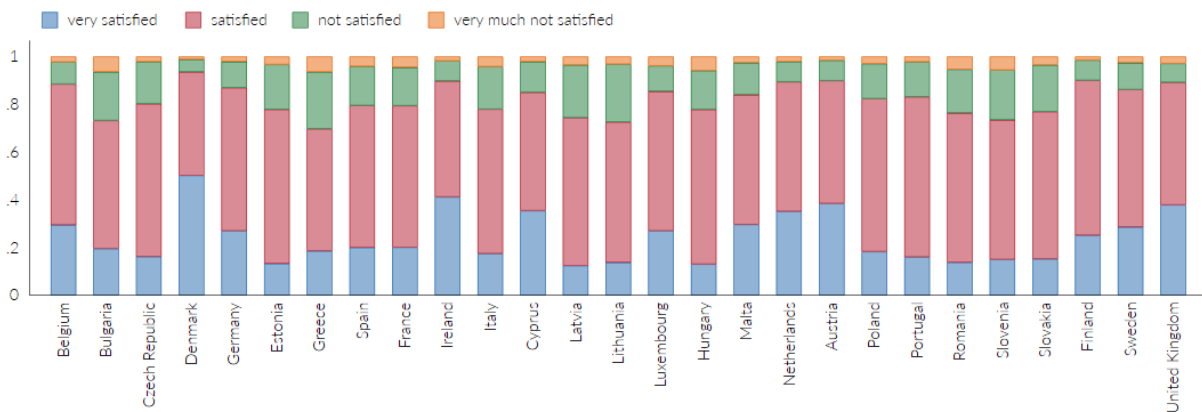
Each wave of the survey provides extensive information about the socio-demographic individual characteristics. In addition to gender and age, individuals report the size of the household, the number and age of children within household, their subjective health status, occupation, and industry. The individuals report also the type of contract (temporary or permanent) and the hours worked (both as a continuous variable and as part-time/full-time dummy). Self-employment is reported separately from wage-employment. Finally, the commute time is reported.

When describing the working conditions, the individuals report a wide battery of answers concerning hazardous conditions. These questions are grouped in three categories: direct hazard, discomfort and wearing protective gear.⁷ . We utilize this battery of questions to construct a

⁶Due to massively missing data for some of the key variables, we cannot use the data for the Balkans, Norway and Switzerland as well as Turkey.

⁷The first list includes: working with vibrations from hand tools, machinery, etc; noise so loud that one needs to raise voice to talk to people, high temperatures which make one perspire even when not working; low temperatures indoors or outdoors; breathing in smoke, fumes, powder or dust; breathing in vapors, handling or being in skin contact with

Figure 1: Job satisfaction across countries in waves in EWCS



Notes The graph is based on EWCS data from all waves.

variable ranging from 0 to 5, where the value of 0 describes no hazardous conditions at work and 5 describes 5 or more hazardous conditions at work.

In addition, the workers describe their working arrangements: working on Saturdays, Sundays, during the nights and long hours. The individuals report how many times a month they had to work in either of those irregular arrangements. We utilize this information to obtain incidence indicators. As of the third wave, the workers report if they work in fixed starting and finishing times.

All but the first wave ask the question *On the whole, are you very satisfied, satisfied, not very satisfied or not at all satisfied with working conditions in your main paid job?*. This variable provides our main indicator of job satisfaction (*JS*). As clear from the question formulation, the answers come on a four-levels Likert scale. On average, 25% of individuals are very satisfied, with additional 57% of individuals being satisfied. The distribution of job satisfaction across countries is portrayed in Figure 1.

The most recent available wave of EWCS data formulated a battery of questions concerning the feelings about one's job and working place. The questions fall under an umbrella *How often you feel this way?* and they include:

- At my work I feel full of wave energy
- I feel exhausted at the end of the work day
- I am enthusiastic about my job
- I doubt the importance of my work
- Time flies when I am working

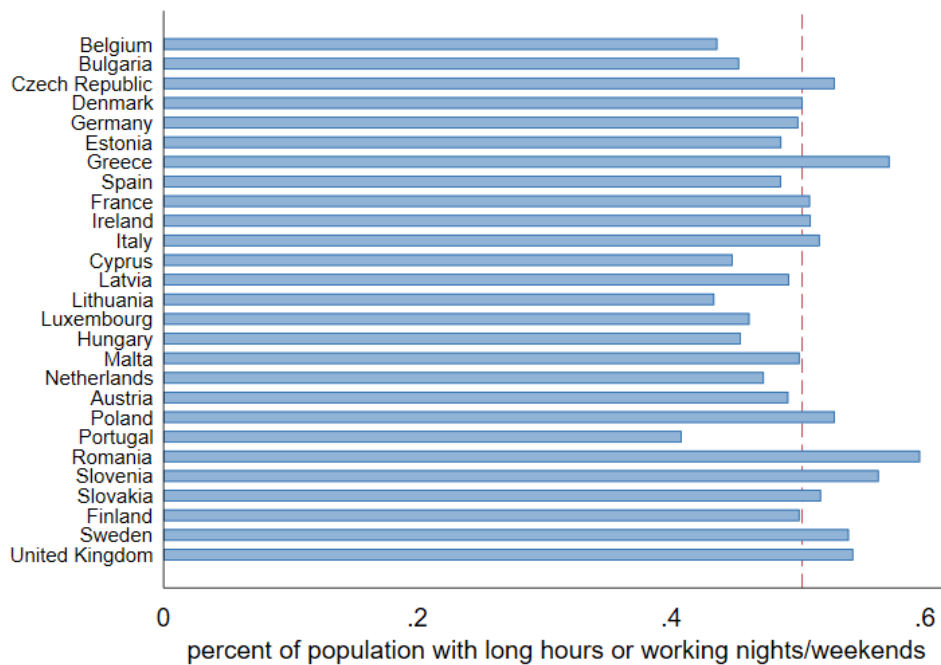
These questions are answered on a five-levels Likert scale ranging from *always* through *most of the time*, *sometimes*, up to *rarely* and *never*. We utilize these variables to construct an index of job satisfaction using factors (*JS_f*). Figure A1 in the Appendices juxtaposes this measure and self-reported job satisfaction (*JS*). The two measures of job satisfaction have a correlation of 0.16 ($p - value = 0.00$). This level of correlation may be indicative that not all of the items from the

chemical products or substances; tobacco smoke from other people; handling or being in direct contact with materials which can be infectious. The second list includes: tiring or painful positions; lifting or moving people; carrying or moving heavy stuff; repetitive hand or arm movements. The third question concerns wearing personal protective equipment.

EWCS questionnaire are actually taken into consideration by the workers, when they report directly their job satisfaction.

First, we construct two synthetic variables, which tag workers in non-standard employment: NSE_1 and NSE_2 . The first one accounts for the features of non-standard employment which facilitate the transgression between work and private life. For all the variables, which are a number in the original data, we construct prevalence dummies: if an event occurs at all, the dummy takes on the value of 1. We then construct a synthetic measure taking on the value of 1 when the worker reports prevalence of either working on Saturdays, working on Sundays, working long hours or working during the nights. The variables needed to construct NSE_1 are available as of third wave of EWCS, and cover the years 2000-2015. On average, in each available wave, roughly 51% of the individuals fall into this category (with a cross-country coefficient of variation = 1.03). We report the distribution of NSE_1 across countries in Figure 2 .

Figure 2: Non-standard employment NSE_1 across countries in EWCS

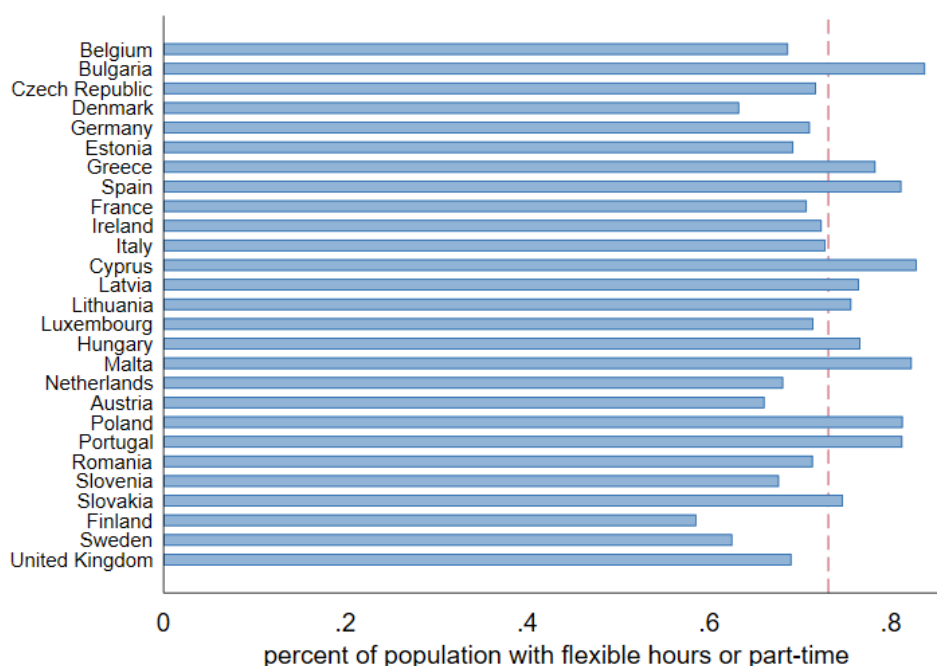


Notes The graph is based on EWCS data from waves 3-6. Non-standard employment facilitating transgression is captured by the prevalence indicators of: working on Saturdays, Sundays, nights or in long hours at least once in a given month. The dashed red line indicates an unweighted average across countries.

Second, we construct a synthetic variable tagging the potential to ease the burden of work obligations on private life, NSE_2 . This indicator takes on the value of 1 for the individuals who report working part-time and working at flexible starts and finishes. Ideally, one would like to be able to assert who specifies the start and finish hours, but this variable is available in the EWCS data only as of the fourth wave. On average, 34% of the individuals report that their working conditions include variable start and finish times or part-time work (with a cross-country coefficient of variation = 0.72). The distribution of NSE_2 across countries is reported in Figure 3 .

Ideally, one would prefer to study each of the above six features separately from the others.

Figure 3: Non-standard employment NSE_2 across countries in EWCS



Notes The graph is based on EWCS data from waves 3-6. Non-standard employment facilitating flexibility is captured by working at flexible starts and finishes (1 if yes, 0 otherwise), and working part-time (1 if yes, 0 otherwise). The variable takes on the value of 1 when either of these conditions is met. The dashed red line indicates an unweighted average across countries.

However, their prevalence is too low to permit viable statistical analysis. Grouping of variables offered above is consistent with our theoretical foundations (Kossek et al., 2004). Specifically, using the individual worker perspective, we group the types of flexible working arrangements which make it harder for the worker to manage the boundaries between work and life (NSE_1) and the types of arrangements which make it easier to manage these boundaries (NSE_2).

Table 1 provides the descriptive statistics of our sample. We report the total sample as well as subsamples split by the status on NSE_1 and NSE_2 . 53% of our sample are women. Women are also more frequent in the case of $NSE_1 = 0$ and $NSE_2 = 1$ subsamples, i.e. females are more likely to work part-time or on flexible schedules and less likely to work on weekends and do overtime. Roughly 14% of individuals in our sample live in single household, and this proportion is fairly similar across subsamples. Approximately 12% of the sample reports having one or more children under the age of 7 in the household and this characteristic too is similar across subsamples. The households in our sample are rarely multi-generational, only 2-3% of the individuals live jointly with an elderly. The fact that these characteristics are so similar across the subsamples suggests that the link between caring obligations and non-standard working arrangements is not very strong in Europe. This observation is consistent with the conflicting empirical evidence presented by (Wheatley, 2017): part-time employment exhibits a complex relationship with caring obligations and professional aspirations.

Table 1 reports also the prevalence of irregular or difficult working arrangements in the total sample and across subsamples. Approximately 35% of the sample works on Saturdays and 20%

Table 1: Descriptive statistics

Variable	Full sample	NSE_1		NSE_2	
		=0	=1	=0	=1
% satisfied with their job	82.4%	85.6%	79.3%	81.3%	82.8%
% of women	53.0%	58.4%	47.7%	43.9%	56.0%
% of single hh	13.8%	13.5%	14.1%	15.5%	13.3%
% of hh with a child aged<7 yo	12.5%	12.3%	12.7%	13.3%	12.2%
% of hh with an elder member	1.8%	1.8%	1.7%	1.5%	1.8%
% working on Saturdays	34.8%	0.0%	69.0%	43.7%	31.8%
% working on Sundays	19.9%	0.0%	39.4%	28.9%	16.8%
% reporting long hours	26.1%	0.0%	51.7%	48.9%	18.4%
% working during nights	16.4%	0.0%	32.5%	25.6%	13.3%
% working part-time	12.3%	14.0%	10.7%	0.0%	16.5%
% working w/ flexible schedules	70.9%	81.5%	60.5%	0.0%	94.8%
% with long commute	28.3%	27.9%	28.6%	33.1%	26.6%
working in hazardous conditions (count)	2.91	2.67	3.15	2.88	2.92
% reporting hours fit their schedules	82.4%	92.0%	72.9%	73.4%	85.4%
% reporting supportive colleagues	92.4%	91.7%	93.1%	93.3%	92.1%
% reporting enough time to finish tasks	92.1%	93.5%	90.7%	90.4%	92.7%
Observations	91 590	45 431	46 159	23 051	68 539

Notes: Table presents EWCS data from wave 3 onward. We report the sample means, weighted. In addition to the reported variables, we use also the information about age (categorized into 5-year age groups), industry (categorized into agriculture, manufacturing, market services, non-market services and others) and occupation (grouping ISCO categories into low-skilled, medium-skilled and high-skilled).

on Sundays. Roughly 26% of workers do overtime and 16% work at nights. Clearly, the figures are much higher for those in $NSE_1 = 1$ subsample (by definition) and those in $NSE_2 = 0$ subsample (workers not working part-time or on flexible schedules are more likely to work on weekends or overtime). Roughly 12% of the sample works part-time and 70% reports flexible schedules. In the case of $NSE_2 = 1$ subsample we observe approximately 16% working part-time and 95% on flexible schedule. The percentages above the average appear also in the case of $NSE_1 = 0$, i.e. workers who do not work on weekends or do overtime are more likely to work part-time or on flexible schedules. In terms of difficult working arrangements we observe a constant across subsample proportion of 30% of workers who report commuting to work time of more than 1 hour per day. Additionally, workers on average report that they are exposed to three hazardous working conditions, such as vibrations, loud noise, high or low temperatures, dust, chemicals etc.; the proportion again is constant across subsamples.

Finally, Table 1 shows workers' subjective opinion on their working environment. Roughly 82% of sample reports that working hours fit their schedules. The proportion is lower for workers in $NSE_1 = 1$ and $NSE_2 = 0$. The fraction of workers who claim that they have supportive colleagues at work or enough time to finish their work tasks is slightly above 90% and similar across subsamples. Most importantly, we observe that approximately 82% of the full sample is satisfied with their job, similarly in the NSE_2 subsamples. However, the proportion differs significantly in the case of NSE_1 subsamples.⁸ The result suggests that workers working on weekends, nights or doing overtime are less satisfied with their job than those with standard working arrangements.

⁸With the value of $t - stat = 25.4$ ($p - value = 0.000$) when testing equality of means in $NSE_1 = 0$ and $NSE_1 = 1$ subsamples.

4 Methodology

We use data from the European Working Conditions Survey. The key challenge in our research is related to the fact that inclination to work in flexible hours is latent and cannot be directly observed. We propose to employ machine learning (ML) methods to uncover the latent link between job satisfaction and the features describing non-standard employment.

The procedure employed proceeds in three steps. First, we split the sample by the measure of flexibility. Second, for each of the groups we obtain a model of job satisfaction *for each of the two subsamples separately* employing machine learning methods. Third, we obtain counterfactual levels of job satisfaction. Specifically, for a group classified as $NSE_i = 1$ we obtain their level of job satisfaction using the model estimated for $NSE_i = 0$ group and vice versa (where $i = 1, 2$). The interpretation of those counterfactual levels of job satisfaction is that this is the level of satisfaction that the individuals of $NSE_i = 1$ would have reported if they were actually $NSE_i = 0$.

Finally, we compare the actual and the counterfactual levels of job satisfaction. We obtain the following potential outcomes:

1. actual and counterfactual job satisfactions remain the same
2. job satisfaction is lower in actual than in counterfactual
3. job satisfaction is higher in actual than in counterfactual

These three potential outcomes have the following interpretation: if job satisfaction for an individual who has $NSE_i = 1$ increases when we apply the model obtained from $NSE_i = 0$, then this person would have been better off working with a standard employment contract. The opposite holds if the individual's job satisfaction declines: that person would be better off working in non-standard working arrangements. Eventually, we can identify four groups: those who work in non-standard arrangements and this is optimal, those who work in standard arrangements and this is optimal, those who work in non-standard arrangements, but would be better off in standard working arrangements and, finally, those who work in standard arrangements and would be better off working in non-standard ones.

The two variables – NSE_1 and NSE_2 – are used independently. Thus, we actually run four independent estimations (for each subsample in each measure). Accordingly, we obtain two counterfactual levels of job satisfaction for each individual.

We utilize machine learning to predict the counterfactual levels of job satisfaction. From the wide range of machine learning classification techniques, the random forest classifier (trade-marked by Breiman, 2001) works best in our data, not only in terms of accuracy (i.e. number of cases correctly predicted), but also in matching the off diagonal entries in the confusion matrix. The algorithm consists of building a large number of individual decision trees, while the final prediction is based on the outcome selected by most trees. An individual tree can be seen as a tool which predicts the value of a target variable (here job satisfaction) based on several input variables (here individual's and job's characteristics). At each step the algorithm chooses an explanatory variable that best (according to Gini impurity measure) splits the data into separate groups. The trees are appropriately pruned by selecting the optimal parameter values for the tree's complexity

and size.

Specifically, for each out of four alternative models we construct in Python 3.8 a random forest (sklearn.ensemble.RandomForestClassifier) with 500 trees on bootstrapped sub-samples of the original data. We set the minimum number of samples essential to split a node to 2, and the minimum number of samples to be at a leaf to 1. The algorithm employs sample weights. To make sure that unique outcomes get higher weights in the classification problem, we adjust the weights to be inversely proportional to frequencies in the data (standard approach suggested in Chen and Breiman (2004)).

The key advantage of deploying the machine learning approach is that the counterfactual job satisfaction is not estimated parametrically. Rather it is assigned based on the individual characteristics through classifying the deep linkages between individual characteristics, and job characteristics. Our approach replicates the approach of Athey et al. (2021). Table 2 reports the original and implied level of job satisfaction for each of the four models in our study. In the remainder of this paper we refer to model JS as counterfactual job satisfaction and to true JS as actual job satisfaction.

Table 2: Internal validity of NSE_1 and NSE_2 models

True JS	Model JS $NSE_1 = 0$				Model JS $NSE_1 = 1$			
	1 Cell %	2 Cell %	3 Cell %	4 Cell %	1 Cell %	2 Cell %	3 Cell %	4 Cell %
very satisfied (1)	18,6	4,6	2,2	1,3	18,0	1,9	1,2	0,5
satisfied (2)	11,5	33,5	9,6	5,6	5,2	47,3	4,7	1,9
dissatisfied (3)	1,1	1,5	7,1	1,1	0,6	1,2	13,2	0,6
very dissatisfied (4)	0,1	0,1	0,2	2,0	0,0	0,1	0,0	3,7
N	14.442	17.789	8.840	4.360	10.790	23.091	9.228	3.050

True JS	Model JS $NSE_2 = 0$				Model JS $NSE_2 = 1$			
	1 Cell %	2 Cell %	3 Cell %	4 Cell %	1 Cell %	2 Cell %	3 Cell %	4 Cell %
very satisfied (1)	17,8	3,6	1,6	0,8	20,8	2,2	1,1	0,4
satisfied (2)	9,8	39,9	7,1	3,4	6,4	45,3	4,4	1,8
dissatisfied (3)	1,0	1,7	9,4	0,8	0,5	0,9	12,2	0,5
very dissatisfied (4)	0,1	0,1	0,1	2,7	0,0	0,0	0,0	3,3
N	19.986	30.564	12.810	5.179	6.248	11.166	4.257	1.380

Notes: **N** denotes the number of observations in a given subsample. Models defined by the subsamples.

The key disadvantage of using random forests is that in the case of singletons, the model cannot learn the patterns of job satisfaction. Singletons occur, if a given composition of characteristics and outcomes exists solely in either training set or in testing set, but not in both. Since the main interest in our paper lies in the counterfactual level of job satisfaction, we abandon the split to training and testing sets and put all the available observations to the training set.

The consequence of putting all observations in the training set is that we cannot obtain out-of-sample predictions to study the quality of the model: we can do in-sample comparison of model with the data, but we cannot do out-of-sample comparisons. This shortcoming can be addressed using the data from the sixth wave: we provide the correlation between the model JS and the actual JS_I , which is an adequate external validity test given that no variable used to construct JS_I

was used in the modeling of JS . The correlations are obtained for each of the four subsamples across NSE arrangements.

We report partial correlation between the original JS and JS_I for each subsample in the first row of Table 3. On average, the correlation between the composite index and the self-reported level of job satisfaction is approximately 0.2-0.25, highly statistically significant. While we report several coefficients (for each subsample across NSE arrangements), we cannot reject the null hypothesis that these coefficients are one and the same in statistical terms. Notably, the confidence intervals around the point estimates in each column substantially overlap with the other three columns.

Table 3: External validity of NSE_1 and NSE_2 models

JS_I	NSE_1		NSE_2	
	=0 (1)	=1 (2)	=0 (3)	=1 (4)
Raw JS	0.199*** (0.0243)	0.211*** (0.0173)	0.245*** (0.0199)	0.206*** (0.0208)
R^2 raw JS	0.041	0.049	0.073	0.041
Model JS	0.167*** (0.0128)	0.133*** (0.0115)	0.193*** (0.0195)	0.151*** (0.0115)
R^2 model JS	0.039	0.045	0.070	0.030
Observations	11,928	11,936	5,648	18,216

Notes: Each model includes country-fixed effects. Data for JS_I is available only in the sixth wave of EWCS. Constant estimated but not reported. Standard errors clustered at the level of country reported in parentheses, *** denotes significance at $p < 0.01$.

The second panel of Table 3 reports the correlation coefficient for the JS implied by our machine learning model. When compared to the direct measure of raw JS , our counterfactual JS from the model performs well. On average, the correlations amount to 0.13-0.19 and the confidence intervals between the coefficients in the first row and in the second row overlap for each column, except column (2). We infer that, in statistical terms, these estimates are not different from each other.

In parallel to the estimates correlations, also the measures of R^2 are similar between the models with raw JS and the models with the JS implied by our models. Note that the dependent variable and the samples are identical between rows in each of the columns, hence the total sum of squares is the same. While comparison of R^2 in this context is meaningful, it also reveals that the self-reported levels of JS have low ability to explain the variation in the composite index of job satisfaction (JS_I). In summary, we establish that the correlations are remarkably similar across both the original and the implied measure of JS . As a byproduct, we also show that self-reported levels of job satisfaction are not very consistent with a composite index, intended to capture the same phenomenon.

5 Results

We describe the results in three substantive parts. First, we refer to Hypothesis 1 and discuss the extent of mismatch between actual and preferred working arrangements across the countries. Second, we refer to Hypothesis 2 and discuss individual level drivers of this mismatch. Finally, we provide an extensive discussion of our findings in the context of the flexibility enactment theory, as well as from the perspective of empirical methodology.

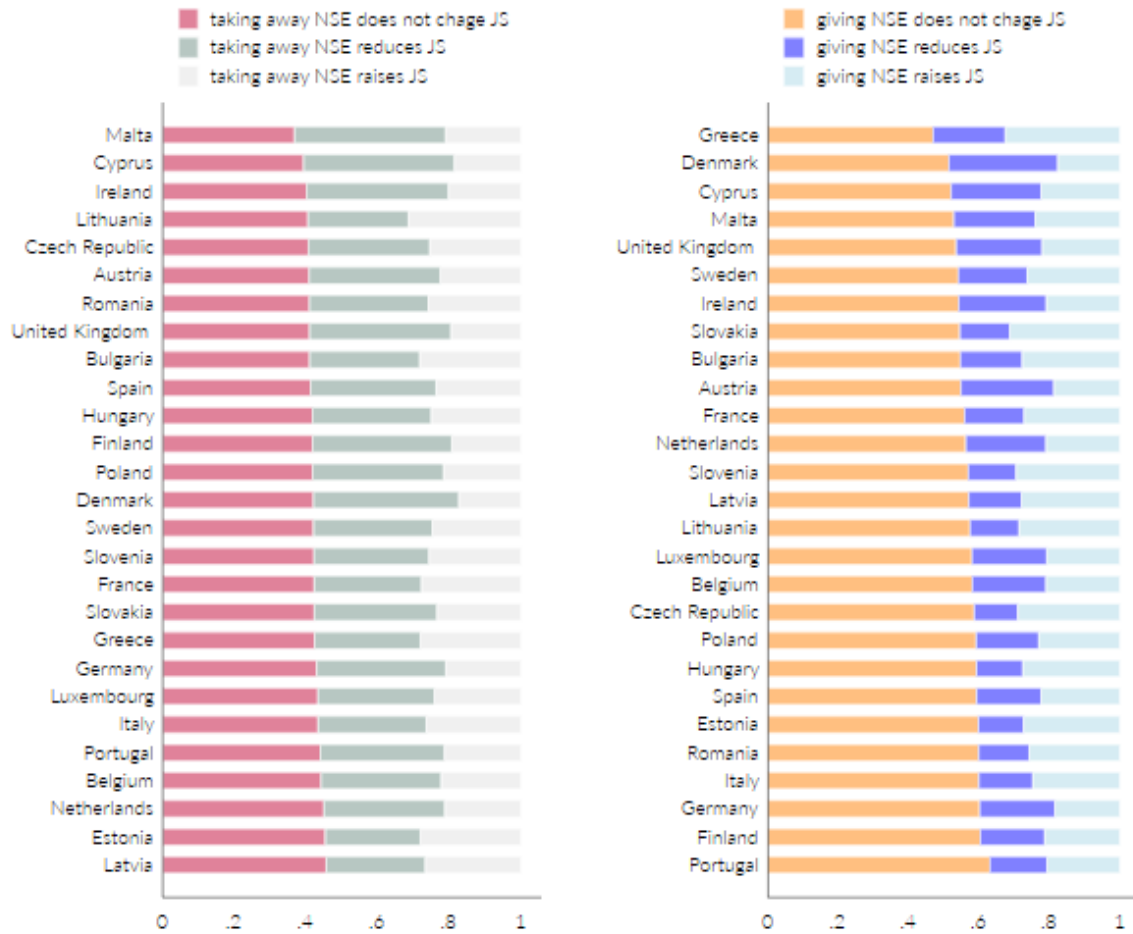
5.1 Mismatch in working arrangements across countries

We verify the first part of Hypothesis 1 by reporting the shares of workers who have nothing to gain or lose from changing the working arrangements and the shares of workers whose job satisfaction could improve or deteriorate subsequent such change. This counterfactual exercise is enabled by the nature of our model: for each worker we can simulate if this person would be better in counterfactual working arrangements, i.e. if this person's job satisfaction would rise or fall, subsequent adjustments in NSE. Figures 4 and 5 report the results of these simulations for NSE_1 and NSE_2 , respectively.

As portrayed by Figure 4, there is little cross-country heterogeneity in the share of individuals whose job satisfaction would remain unchanged if NSE_1 was replaced by regular job arrangements (roughly 40%) or the opposite: if regular arrangements were replaced by NSE_1 (roughly 55%). There is, however, substantial heterogeneity in the share of individuals whose job satisfaction would decline if NSE_1 was taken away (between 15% for the Baltic states and as much as 40% for Denmark as well as Cyprus and Malta). Likewise, generally in Central and Eastern Europe not many workers would suffer if their regular arrangements were replaced by NSE_1 (roughly 10%), but such share of workers exceeds 20% in BeNeLux, Germany, Austria and Denmark. The same ordering of countries applies to raising job satisfaction from being given the opportunity to work with NSE_1 , with the difference that now the ordering is the opposite (BeNeLux, Germany and Austria would gain the least, whereas CEECs gain the most). Finally, merely 10% of workers would experience decline in job satisfaction if NSE_1 was taken away from them in Scandinavian countries or the UK, whereas this fraction would exceed 20% in France, Greece, Italy, Bulgaria and the Baltic states. This last result is particularly interesting given that the features of working arrangements which fall under NSE_1 are typically thought of as facilitating transgressions from work to private life. Meanwhile, our results seem to suggest that many individuals are able to individually manage the boundaries between these two spheres and their job satisfaction would not be diminished by the necessity to actively engage in setting boundaries.

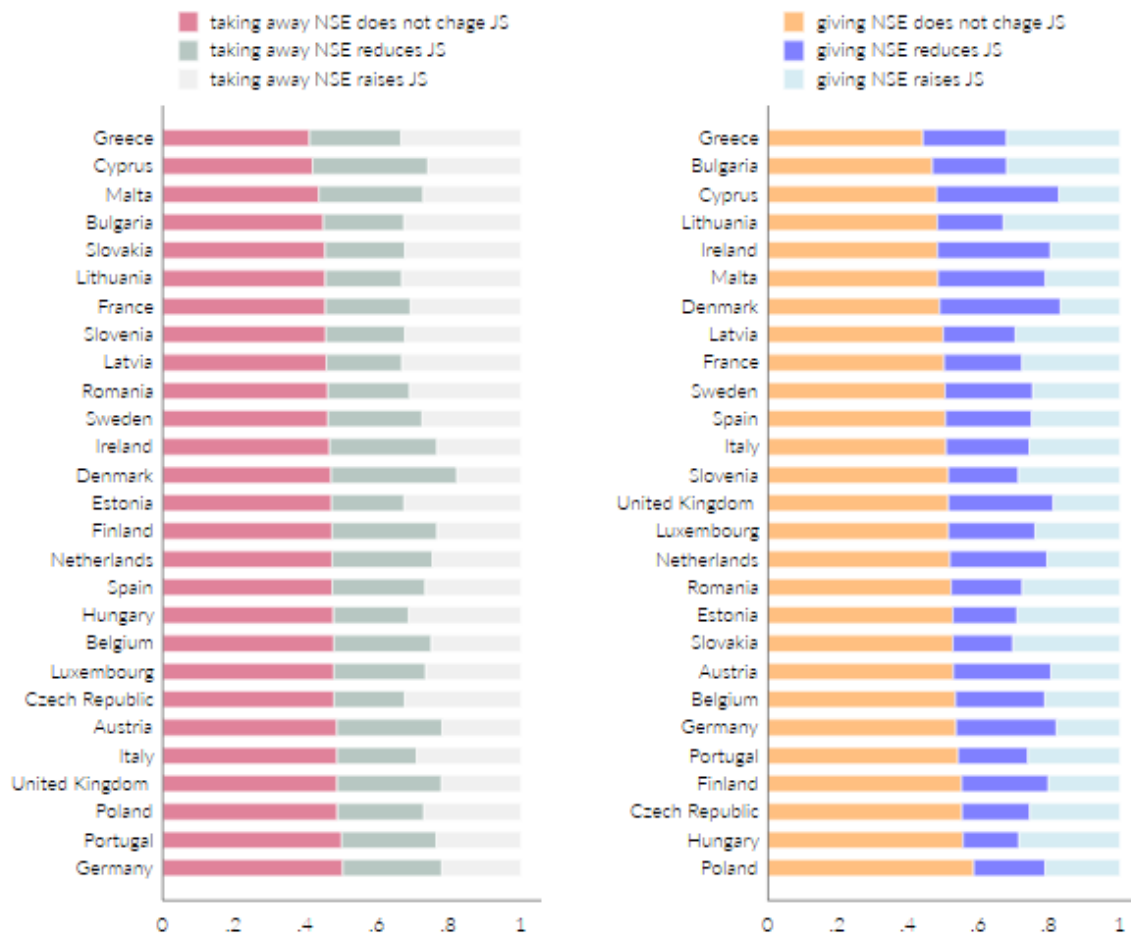
Cross-country heterogeneity is even more paramount in the case of NSE_2 . Roughly half of the workers would experience no change in job satisfaction if NSE_2 was replaced by regular arrangements and vice versa, that is if they were given the opportunity to work in those arrangements. This share is rather similar across countries, as portrayed in Figure 5. However, there is substantial cross-country heterogeneity in the share of winners (raised job satisfaction) and losers (reduced job satisfaction). When it comes to taking NSE_2 away or giving this opportunity to workers, the

Figure 4: Changes in job satisfaction if mismatch was reduced for NSE_1



Notes Taking away NSE refers to a counterfactual change in job satisfaction where a person is actually in a job with NSE_1 and is assumed to change job features to no NSE_1 . Giving NS refers to a counterfactual change in job satisfaction where a person is in a actually in a job without NSE_1 and is assumed to change job features to have it. Shares of mismatched and matched individuals obtained with population weights.

Figure 5: Changes in job satisfaction if mismatch was reduced for NSE_2



Notes Taking away NSE refers to a counterfactual change in job satisfaction where a person who currently has NSE_2 is assumed to have no NSE_2 . Giving NSE refers to a counterfactual change in job satisfaction where a person who actually does not have NSE_2 is assumed to have it. Shares of mismatched and matched individuals obtained with population weights.

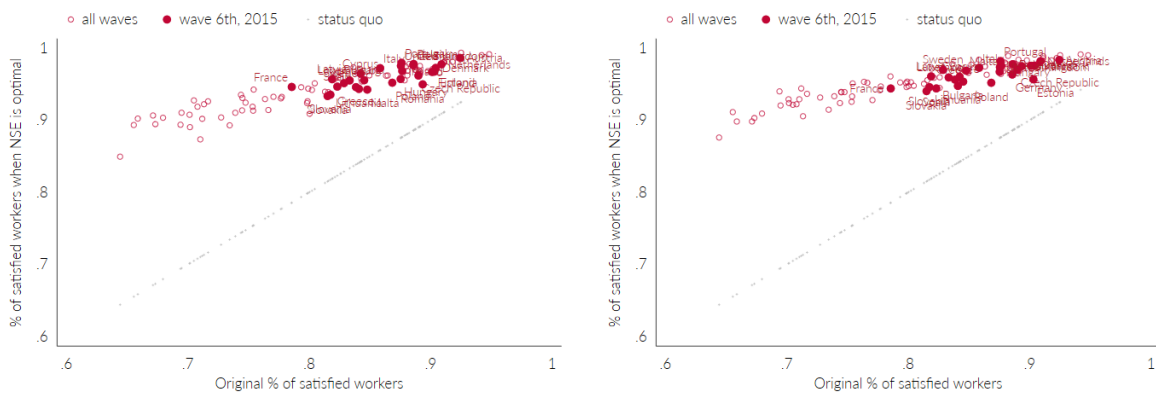
largest share of winning workers would be observed in Central and Eastern Europe and the lowest in Northern Europe. This suggests that the mismatch is the highest in CEECs and the lowest in Denmark, Germany and Austria. Indeed, the share of losers from the change would be the lowest in CEECs and the highest in Northern Europe (with Cyprus and Malta as exception).

Given the substantial differences in prevalence of non-standard employment across these groups of countries, as portrayed in Figures 2 and 3, this counterfactual exercise reveals that the room for improving job satisfaction differs largely across countries. Note that the share of winners and losers from changing the working arrangements in the case of either NSE_1 or NSE_2 is not simply correlated with the prevalence. Indeed, it appears that in some of the countries there is simply too little non-standard employment, while in others the problem is that it is not the “right” workers who work in regular arrangements. For example, in the group of CEECs, there are countries with both the lowest and the highest prevalence of NSE_1 in the whole sample. Similarly non-schematic is prevalence of NSE_2 , frequent especially in Southern Europe, but also in Poland.

The relationship between prevalence of non-standard employment and the share of workers who would gain or lose from changing their working arrangements is non-monotonous.

In order to verify the aggregate part of Hypothesis 1, we utilize the counterfactual levels of job satisfaction obtained through models and assign each individual a maximum value of three options: the factual job satisfaction JS and the two model implied levels (if non-standard working arrangements are taken away from those who have them and if they are given to those who do not have them). This way, we obtain for each individual in our sample “the best of all worlds”. For each country and wave of the EWCS data we compute the aggregate share of satisfied individuals, lumping together those who have reported or would have reported “satisfied” or “very satisfied”. The results of this exercise are reported in Figure 6. For illustrative purposes we emphasize the most recent available wave of EWCS and label countries, the other waves are reported as less discernible scatters. Status quo (the 45 degree line) is displayed in gray. Instead of a line, we use the scatters to visualize the relative density of the factual JS values (% of satisfied individuals) across countries and waves.

Figure 6: Aggregate implied job satisfaction: NSE_1 (left) and NSE_2 (right)



Notes Optimal change in NSE’s involves assigning each individual such NSE that his/her job satisfaction is maximized among three options: factual JS and model implied JS either taking away flexibility (from those who have it) or giving it (to those who do not have it). For illustrative purposes, we report shares of satisfied individuals (taking satisfied and very satisfied as one group). Shares of mismatched and matched individuals obtained with population weights.

In all countries and waves the share of satisfied individuals is increased, as one would expect, given the nature of our counterfactual exercise. However, the key contribution of this graph is the quantification of how many workers in each country could be lifted from dissatisfaction to satisfaction by the change in the working flexibility. We find that this is a sizable effect, ranging between roughly 14-19%. Note that this vast improvement in job satisfaction does not necessitate the rise of the use of NSE. Clearly, raising the use of NSE could be potentially detrimental to businesses and labor productivity. However, as evidenced by Figures 4 and 5, a large part of this gain could be achieved by workers reallocating themselves between jobs with and without non-standard arrangements. The persistence of this mismatch across countries and waves speaks of welfare costs stemming (partially) from labor market frictions preventing a better sorting of workers across working time arrangements.

5.2 Individual drivers of working arrangements mismatch

In order to verify Hypothesis 2 we run a set of ordered logistic regressions for various cases of the mismatch (taking away or giving NSE_1 and NSE_2) with distinct sets of regressors. The dependent variable is the change in job satisfaction that workers would see from switching towards a counterfactual working arrangement. It is the difference between the imputed counterfactual model and the actual job satisfaction. We can write this difference as:

$$\Delta JS_{i,NSE} = \text{True } JS_i - \text{Model } JS_{i,NSE}$$

where $\Delta JS_{i,NSE}$ is the change in job satisfaction, and $\text{Model } JS_{i,NSE}$ indicate the levels for worker i in $NSE \in \{NSE_1, NSE_2\}$. For example, imagine an individual who works part-time ($NSE_2 = 1$) and reports low job satisfaction (4). Given this individual's personal and job's characteristics, our model indicates that had this worker been employed in a full-time contract (i.e. taking away NSE_2), their job satisfaction would have been 1. Then, the value of the dependent variable in the ordered regression equals the change ($\Delta JS_{i,NSE_2} = 4 - 1 = 3$). Given that job satisfaction variables have four levels, the differences take on seven different values (three positive, three negative and no change). The bigger values of $\Delta JS_{i,NSE}$ correspond to larger improvement in job satisfaction in case counterfactual working arrangements were to become factual. Analogously, negative values of $\Delta JS_{i,NSE}$ imply a decline in job satisfaction.

We focus on exploring the role of gender and child-caring in the mismatch of the working arrangements. We thus include the personal characteristics of the main importance, i.e. gender and presence of child under 7 years of age in the household. We also adjust for individual age. Ideally, we would like to be able to adjust the estimates for subjective perception of the work related issues such as ability to make ends meet, worry about work when not working, family work conflict, workplace where employer sets working time and enough time to finish tasks. However, these variables are only available as of the fifth wave. We thus provide two sets of estimates: models for the full available sample and basic regressors (age, gender, children in the household) and models for the full set of regressors, respectively in Tables 4 and 5. All specifications include country fixed-effects and year fixed-effects. If we were to report marginal effects, given the seven levels of our dependent variable, there would be seven effects to report for each explanatory variable. However, our explanatory variables take only two values: 0 and 1. Hence, the coefficients report an average of marginal effects weighted by the frequency of each level of dependent variable and thus they have a clear interpretation (though, naturally, qualitative rather than quantitative). We report marginal effects for two key variables graphically in Figures A2 and A3 in the Appendix.

Gender Across all specifications in both Table 4 and Table 5 we observe large, positive and highly significant coefficients for women. This suggests that the mismatch is systematically stronger for women than it is for men. Moreover, we may even infer that women have more to gain from changing their working arrangements than men do. Indeed, women appear to be more frequently mismatched in terms of non-standard employment forms and change in work arrangements is (on

average) beneficial for them. The highest estimated coefficient (in terms of magnitude) is observed in the giving NSE_2 scenario, which means that women benefit the most when they are offered to change into part-time or flexible schedule regimes. The exact interpretation of the estimated coefficients is that the ordered log-odds for a (childless) woman being in a higher job satisfaction category is 0.06-0.18 more than equivalent (childless) man. The models can also be interpreted in terms of the marginal effects, which are presented in Figure A2 in the Appendix.

Caring obligations Table 4 reveals that being a father of children under 7 years of age matters a lot for changes in job satisfaction in the case of NSE_1 . The full sample and narrow set of control factors in Table 4 suggests that the interaction between gender and having young children is not statistically significant, meaning that the impact of children is the same for both genders. Reducing NSE_1 improves parents' job satisfaction and giving NSE_1 to parents on standard contracts is likely to reduce it. These results imply that parents may be particularly often experiencing the risk of work-life transgressing into their private life. Intuitively, parents of young children working on the weekends, at nights or doing overtime, would be better off when moved to standard work arrangements. The estimated coefficient in giving NSE_1 model is higher (in magnitude) suggesting that the disadvantage from moving to irregular work regime is larger (more likely) than the benefit in the reverse scenario. The issue does not need to be related to working hours per se: the effects for NSE_2 are smaller and less precisely estimated. Admittedly, precision appears highly important when we compare the estimates from Tables 4 and 5: in more recent sample, taking away flexibility appears to have similar effects for both NSE_1 and NSE_2 . However, the estimate of giving both NSE_1 and NSE_2 is much lower and less precisely estimated in Table 5.

The recent waves of EWCS reveal a less straightforward image. Table 5 reveals that giving NSE_1 has large negative implications for job satisfaction for mothers, but not for fathers, i.e. only mothers are expected to decrease their job satisfaction when moving to non-standard working time arrangements. Note that the estimated coefficient on the interaction is higher (in absolute terms) than the coefficient on gender, overtaking the positive effect. Detailed marginal effects of children under 7 years old on the probability of a change in the job satisfaction in the counterfactual NSE's are presented in Figure A3 in the Appendix.

The existing literature (Bainbridge and Townsend, 2020) argues that parents (or, more generally, care providers) appreciate flexibility. Our results refine this conjecture in a sense that there are some forms of non standard working arrangements that can be helpful to some parents. However, the relationship is far from universal for both the parents and the forms of working arrangements. Intuitively, some individuals have greater capacity to manage e.g. irregular hours – whereas for some it is the fixed hours that helps to successfully combine family and professional life.

Other factors Table 5 summarizes the impact of other variables of interest on the mismatch between actual and preferred work flexibility. We start by analyzing the indicator for the ability to make ends meet. We might treat it as a proxy for a income satisfaction, or at least a personal perception of income. Interestingly, we observe a negative, significant coefficient in all models.

This means workers who consider their income satisfactory are less likely to be employed in the working arrangements different than their preferred (optimal) ones.

Similar results are obtained for worrying about work and sufficient time to complete tasks: people who report sense of security associated with their work are less likely to be working in mismatched arrangements, even if they worry about work “after hours”. In other words, their job satisfaction cannot be raised by changing NSE’s. It appears that sense of insecurity at work attenuates the negative effects of any mismatch and makes workers perceive their working arrangements as suitable.

Whereas the income and the insecurity at work appear to have large effects, the family-work conflict displays small effects: they are negative and statistically significant (except giving NSE_2). We interpret this finding as an indication that individuals who experience tighter clash between professional and private life are less likely to benefit from any change. This result is particularly relevant from the perspective of boundary management theory and flexibility enactment theory: the family-work conflict is not related to working arrangements per se (or at least not to a large extent). Rather it may be the character of work and its demands that impose the sense of work-family conflict.

EWCS provides also information on the discretion to set the specific working hours. This variable is only available since the third wave and thus could not be used for the machine learning algorithms, but in Table 5 we may observe its correlation with the mismatch. It appears that the discretion over own working time is irrelevant for changes in job satisfaction, with the exception of giving NSE_2 . Intuitively, job satisfaction is likely to increase for workers whose *employer sets the working hours* when they are able to claim discretion over their starting and finishing times. That is, workers in standard employment arrangements without discretion to set their own hours would see higher job satisfaction levels in part-time, flexible schedules.

Table 4: Impact of gender and children on job satisfaction when changing work flexibility regime

	NSE_1		NSE_2	
	taking away	giving	taking away	giving
woman	0.17*** (0.03)	0.11*** (0.03)	0.09*** (0.02)	0.18*** (0.03)
child aged<7 yo	0.23*** (0.03)	-0.19*** (0.04)	0.09*** (0.03)	-0.07 (0.07)
woman × child aged<7 yo	-0.07 (0.05)	-0.01 (0.05)	0.04 (0.04)	0.08 (0.08)
Observations	46159	45431	68539	23051

Notes: Table presents the results of estimating an impact of gender and children on workers job satisfaction when changing (taking away or giving) work flexibility regime. Reported are point estimates. The dependent variable is the change in job satisfaction index (ΔJS_T). Models include age and age squared, individual countries indicators and wave indicators as regressors, but no other controls. Standard errors presented in parentheses. ***, ** and * denote significance at $p < 0.01$, $p < 0.05$, and $p < 0.1$, respectively.

Table 5: Impact of individual characteristics on job satisfaction when changing work flexibility regime

	<i>NSE</i> ₁		<i>NSE</i> ₂	
	taking away	giving	taking away	giving
woman	0.15*** (0.05)	0.13*** (0.04)	0.07* (0.04)	0.15** (0.07)
child aged<7 yo	0.17*** (0.06)	-0.10 (0.09)	0.15*** (0.04)	-0.03 (0.10)
woman × child aged<7 yo	-0.13 (0.09)	-0.17** (0.08)	-0.09 (0.06)	-0.18 (0.15)
able to make ends meet	-0.31*** (0.05)	-0.44*** (0.05)	-0.22*** (0.03)	-0.51*** (0.06)
enough time to finish tasks	-0.51*** (0.13)	-0.47*** (0.17)	-0.31*** (0.12)	-0.46*** (0.17)
worry about work when not working	-0.27*** (0.05)	-0.37*** (0.12)	-0.31*** (0.07)	-0.31*** (0.09)
family work conflict	-0.03** (0.01)	-0.07*** (0.02)	-0.04*** (0.01)	-0.01 (0.02)
employer sets time	-0.01 (0.04)	0.09 (0.06)	-0.00 (0.04)	0.22*** (0.05)
Observations	11745	11764	17955	5554

Notes: Table presents the results of estimating an impact of selected individual characteristics on workers job satisfaction when changing (taking away or giving) work flexibility regime. Reported are point estimates. The dependent variable is the change in job satisfaction index (ΔJS_I). Models include age and age squared, individual countries indicators and wave indicators as regressors. Standard errors presented in parentheses. ***, ** and * denote significance at $p < 0.01$, $p < 0.05$, and $p < 0.1$, respectively.

5.3 Discussion

Our results show that women are substantially more likely to experience mismatch between their preferred and actual working arrangements. In fact, any change of working arrangements is likely to improve job satisfaction among women. This strong gender gradient appears despite the observation that across all countries, all models and all data subsamples, majority of individuals would experience no change in job satisfaction subsequent a change in work arrangements. Parenting is also relevant: parents appear to gain job satisfaction from moving towards standard working arrangements and lose from more non-standard arrangements. We do not find differential effects for mothers and fathers (except in the recent sample, when giving *NSE*₁).

Our results serve to substantiate the flexibility enactment theory. It appears that some individuals are perfectly capable of actively managing boundaries between work and private life in a sense that the potential non-standard working arrangements do not reduce their job satisfaction. However, there is a group of individuals, whose ability to manage boundaries is insufficient to accommodate the requirements of non-standard arrangements. Their job satisfaction suffers (likely, though unavailable in EWCS, their life satisfaction may suffer as well). Overall job satisfaction can be substantially improved if the mismatch between individual abilities to manage NSE and need to manage NSE is reduced. The policy debate about non-standard arrangements ought to account for the individual inclinations, rather than force one-size-fits-all regulations. Notably, workers' rights protection becomes much more subtle under flexibility enactment theory.

Our results are derived from an ordered logit based on differences between two ordinal vari-

ables. It is likely that the proportional odds assumption may be violated for some of the variables. Formal testing of that hypothesis reveals that the problems appear for some of the control variables: age and country fixed effects. We re-estimated our model re-coding the outcome variable to three levels: decline in job satisfaction (-1), no change in job satisfaction (0) and increase in job satisfaction (1). The key results remain the same and Brant's test no longer rejects the null hypothesis that odds are proportional. The same applies to a model with five levels (when the values of -3 and +3 are top coded to -2 and +2, respectively). We prefer the specification with the broad range of values for the dependent variable, because the change of JS by a factor +2 (-2) has a clear interpretation: that individual has for sure changed from dissatisfied to satisfied (from satisfied to dissatisfied).⁹

One potential concern related to our results stems from the fact that NSE is operationalized through the combination of different forms of working arrangements. Ideally, one would prefer to study each form in isolation from the others, and then gradually build the indicators into more complex ones. With the data limitations of EWCS such analysis would not be feasible, but with other data the step wise analysis could reveal deeper and more refined policy implications. In principle, our method can be applied the same way to individual level data from one specific country. We have experimented with German SOEP data, but the questions related to working arrangement were asked once every few years, which reduces substantially the number of observations with nonstandard employment.

Another potential concern **refers to** the nature of the phenomenon studied in this paper. Admittedly, the inclination to work in non-standard arrangements is unobservable to us, but it is observable to each individual in EWCS. Given this limitation, we infer optimality from patterns from *observationally equivalent* individuals. Based on this premise, we build a model which predicts that a given individual could gain or lose from a change in their working arrangements (being given NSE or having non-standard arrangements replaced by standard ones). This gain/loss is our indication of mismatch. However, it is possible that individuals, observing their own inclinations, purposefully select between standard and non-standard arrangements. In this case, our model would wrongly attribute differences in preferences to a mismatch in working arrangements. However, unobservable (intangible) character of inclination to work in non-standard employment cannot explain away our results, because the mismatches appear to be systematic, whereas there is no reason for them to be systematic across individual characteristics. Specifically, if our results were all about unobservability of inclination, one would have to believe that self-selection to NSE based on inclination is systematically positive among women, positive for taking away both NSE_1 and NSE_2 among parents, etc. This is highly improbable. Thus, while our results ought to be interpreted with caution, they cannot be explained away with unobservables.

Finally, our results would benefit greatly from a richer measure of job satisfaction. Clearly, the two measures available in EWCS are weakly correlated. The composite measure based on specific marker questions do not predict the self-reported levels of job satisfaction. Indeed, vast majority of people reports being satisfied at their job *despite* negative answers to the marker questions. More refined measurement of job satisfaction could permit a more complex machine learning

⁹Results from those additional estimations are available upon request.

classification than currently feasible, potentially revealing refined links between the job features, working arrangements and job satisfaction.

6 Conclusions

In this paper we build on flexibility enactment theory. We leverage this approach to formulate novel research hypotheses about the prevalence of mismatch between the preferred and actual working arrangements. Our approach helps to reconcile two superficially opposite views in the policy debate. The first view emphasizes that flexibility at work is what helps the workers to reach their potential. The second view raises that work tends to transgress to private life, this transgression is easier when the working arrangements leave room for abuse by the demands of the job. The flexibility enactment theory posits that some workers are not immune to such abuse, because they are able to actively manage boundaries and separate work from private life. This theory builds on the role theory and social identity theory as well as spillover theory and boundary theory. Jointly, these theories delineate that the ability to actively manage boundaries is not merely a skill, but actually an ability to derive satisfaction from a job.

We test implications from flexibility enactment theory using data from the European Working Conditions Survey. We deploy modern machine learning methods to uncover (potentially complex and nonlinear) relationships between individual characteristics, family characteristics, job features and job satisfaction, separately for individuals working in standard employment arrangements and in non-standard employment arrangements. These two machine learning models are then used to obtain counterfactual levels of job satisfaction: if an individual working under standard arrangements had been working in non-standard ones and vice versa. These counterfactual simulations permit to identify who would benefit from a change in working arrangements, who would lose, and who is indifferent between the two arrangements. Workers in the first group are denoted as mismatched and those in the latter group as matched.

We characterize the extent and drivers of mismatch among 27 European countries. We show that across countries, there is substantial room for improvement in job satisfaction from simply reallocating the workers across the working conditions. We also show that the mismatch is much more prevalent among women and among parents (we find no additional effect for mothers). We also show that mismatch is less likely to occur among individuals whose income is satisfactory (by their subjective judgment). The mismatch is also less frequent among individuals who believe that they have enough time to complete their tasks. Finally, mismatch is *lower* among individuals who report worrying about work outside their working hours.

Our results have important implications in terms of policy. Given substantial heterogeneity across countries as well as systematic character of mismatch across individuals, no one-size-fits-all policy is likely to be successful in raising the job satisfaction of workers. Whether the use of non-standard working arrangements is curtailed or further popularized, no universality applies. Indeed, the only potentially universal solutions should focus on introducing the boundary management skills into educational curricula. Otherwise, individuals are highly heterogeneous in

their perception of suitable working arrangements. Furthermore, the workers' inclination to engage in non-standard arrangements is likely to change over the life-cycle and depends on personal circumstances.

Finally, our results call for further theorizing on the links between job satisfaction and working conditions. While the flexibility enactment theory appears to find strong confirmation in the data, there seem to be several particularly promising areas for exploration. First, it appears that the notion of social identity may be more directly intertwined with the flexibility enactment. Specifically, the identity may be an important pillar of the boundary management strategy, which has been evidenced by the existing literature on e.g. ethnicity-specific work ethics. While it was demonstrated that managing spillovers may be viewed as a distinguishing characteristic of some groups, it is not warranted that the extent of satisfaction derived from this ability is automatically related to ability itself. In other words, it may be that one self-identifies with a group which manages spillovers well and engages non-standard working arrangements to extract rents related to this skill, but one does not derive satisfaction from these arrangements. Feeling competent to manage some situations is not automatically equivalent to an appreciation of the situations themselves.

Second, it appears that the existing theories expect the non-standard working arrangements to facilitate transgressions from work to private lives (hence the need to *manage* the spillovers). There exists large literature on transgressions in occupations with long hours (medical professions, police force). However, to some individuals there may be intrinsic value to arrangements to their occupations. Indeed, in many professional groups the ability to work at all necessitates non-standard arrangements: night hours and long commutes on some days coupled with large extent of discretion over working time in other days (e.g. actors, musicians, tv and radio producers), short hours combined with lack of discretion over which specific hours to work (e.g. teachers). Better understanding of intrinsic value of the non-standard arrangements through the lenses of flexibility enactment theory permit would help us refine the meaning of the term "transgression" and thus the mechanisms of job (and life) satisfaction.

Last, but not least, the gender context of the flexibility enactment theory is needed. Our results show that women are much more likely to experience mismatch in their working arrangements, but we find little support that this effect is driven by parenting. Indeed, for most types of mismatches, parents of both genders are equally likely to benefit from a change in arrangements. Given these findings, one cannot explain persistent and large gender differences in mismatches with motherhood burdens associated with childbearing and rearing. These large gender differences call for more thorough exploration into social identity theory and boundary theory.

References

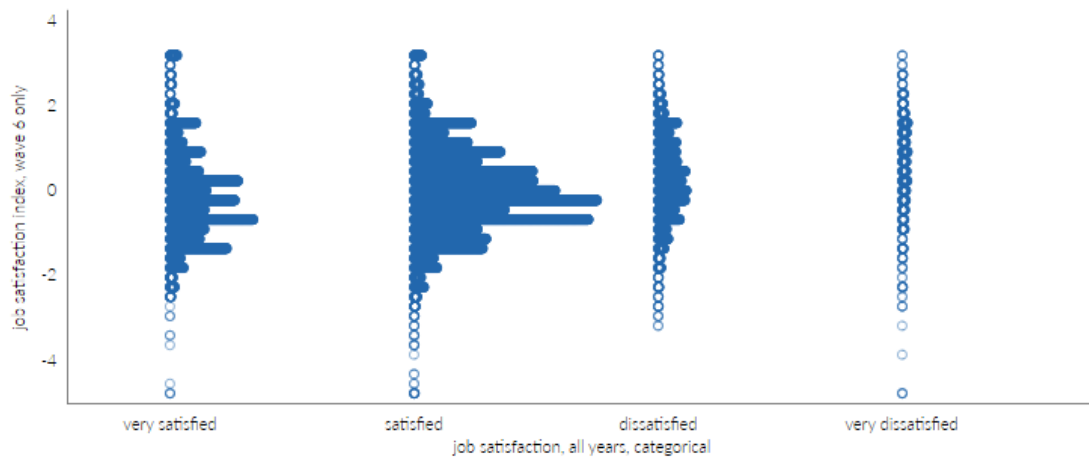
- Albanesi, S. and Olivetti, C. (2009). Home production, market production and the gender wage gap: Incentives and expectations. *Review of Economic Dynamics*, 12(1):80–107.
- Athey, S., Bayati, M., Doudchenko, N., Imbens, G., and Khosravi, K. (2021). Matrix completion methods for causal panel data models. *Journal of the American Statistical Association*, pages 1–15.
- Atkinson, C., Lucas, R., and Hall, L. (2011). Flexible working and happiness in the nhs. *Employee Relations*, 33:88–105.
- Bainbridge, H. T. and Townsend, K. (2020). The effects of offering flexible work practices to employees with unpaid caregiving responsibilities for elderly or disabled family members. *Human Resource Management*, 59(5):483–495.
- Barney, C. E. and Elias, S. M. (2010). Flex-time as a moderator of the job stress-work motivation relationship: A three nation investigation. *Personnel Review*.
- Bellmann, L. and Hübler, O. (2020). Working from home, job satisfaction and work–life balance—robust or heterogeneous links? *International Journal of Manpower*.
- 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*, 130(1):165–218.
- Breiman, L. (2001). Random forests. *Machine Learning*, 45(1):5–32.
- Brescoll, V. L., Glass, J., and Sedlovskaya, A. (2013). Ask and ye shall receive? the dynamics of employer-provided flexible work options and the need for public policy. *Journal of Social Issues*, 69(2):367–388.
- Cech, E. A. and Blair-Loy, M. (2014). Consequences of flexibility stigma among academic scientists and engineers. *Work and Occupations*, 41(1):86–110.
- Cha, Y. and Weeden, K. A. (2014). Overwork and the slow convergence in the gender gap in wages. *American Sociological Review*, 79(3):457–484.
- Chen, C. and Breiman, L. (2004). Using random forest to learn imbalanced data. *University of California, Berkeley*.
- Clark, S. C. (2000). Work/family border theory: A new theory of work/family balance. *Human relations*, 53(6):747–770.
- Collewet, M., de Grip, A., and de Koning, J. (2017). Conspicuous work: peer working time, labour supply, and happiness. *Journal of Behavioral and Experimental Economics*, 68:79–90.
- Cortes, P. and Pan, J. (2019). When Time Binds: Substitutes for Household Production, Returns to Working Long Hours, and the Skilled Gender Wage Gap. *Journal of Labor Economics*, 37(2):351–398.

- Cubas, G., Juhn, C., and Silos, P. (2019). Coordinated work schedules and the gender wage gap. NBER WP 26548, National Bureau of Economic Research.
- Duchini, E. and van Effenterre, C. (2018). School schedule and the gender pay gap. Policy Brief 26, Institute for Public Policy.
- Dumas, T. L. and Sanchez-Burks, J. (2015). The professional, the personal, and the ideal worker: Pressures and objectives shaping the boundary between life domains. *The Academy of Management Annals*, 9(1):803–843.
- Flabbi, L. and Moro, A. (2012). The effect of job flexibility on female labor market outcomes: Estimates from a search and bargaining model. *Journal of Econometrics*, 168(1):81–95.
- Hamplová, D. (2019). Does work make mothers happy? *Journal of Happiness Studies*, 20(2):471–497.
- Hayman, J. R. (2009). Flexible work arrangements: Exploring the linkages between perceived usability of flexible work schedules and work/life balance. *Community, work & family*, 12(3):327–338.
- Kahn, R. L., Wolfe, D. M., Quinn, R. P., Snoek, J. D., and Rosenthal, R. A. (1964). *Organizational stress: Studies in role conflict and ambiguity*.
- Katz, D. and Kahn, R. L. (1978). *The social psychology of organizations*, volume 2. Wiley New York.
- Kossek, E. E., Lautsch, B. A., and Eaton, S. C. (2004). Chapter 12: Flexibility enactment theory: Implications of flexibility type, control, and boundary management for work-family effectiveness. In *Work and Life Integration: Organizational, Cultural and Individual Perspectives*. Psychology Press.
- Lee, M. D., MacDermid, S. M., Williams, M. L., Buck, M. L., and Leiba-O’Sullivan, S. (2002). Contextual factors in the success of reduced-load work arrangements among managers and professionals. *Human Resource Management*, 41(2):209–223.
- Lott, Y. (2015). Working-time flexibility and autonomy: A european perspective on time adequacy. *European Journal of Industrial Relations*, 21(3):259–274.
- Lott, Y. (2020). Does flexibility help employees switch off from work? flexible working-time arrangements and cognitive work-to-home spillover for women and men in germany. *Social Indicators Research*, 151(2):471–494.
- Mas, A. and Pallais, A. (2017). Valuing alternative work arrangements. *American Economic Review*, 107(12):3722–3759.
- McCarthy, A., Cleveland, J. N., Hunter, S., Darcy, C., and Grady, G. (2013). Employee work-life balance outcomes in ireland: a multilevel investigation of supervisory support and perceived organizational support. *The International Journal of Human Resource Management*, 24(6):1257–1276.

- Munsch, C. L. (2016). Flexible work, flexible penalties: The effect of gender, childcare, and type of request on the flexibility bias. *Social Forces*, 94(4):1567–1591.
- Nippert-Eng, C. (1996). Calendars and keys: The classification of “home” and “work”. 11(3):563–582.
- O’Connor, L. T. and Cech, E. A. (2018). Not just a mothers’ problem: The consequences of perceived workplace flexibility bias for all workers. *Sociological Perspectives*, 61(5):808–829.
- Rudman, L. A. and Mescher, K. (2013). Penalizing men who request a family leave: Is flexibility stigma a femininity stigma? *Journal of Social Issues*, 69(2):322–340.
- Schieman, S. and Glavin, P. (2016). The pressure-status nexus and blurred work–family boundaries. *Work and Occupations*, 43(1):3–37.
- Secunda, P. M. (2019). The employee right to disconnect. *Notre Dame Journal of International Comparative Law*, 9:1.
- Vandello, J. A., Hettinger, V. E., Bosson, J. K., and Siddiqi, J. (2013). When equal isn’t really equal: The masculine dilemma of seeking work flexibility. *Journal of Social Issues*, 69(2):303–321.
- Westman, M. and Piotrkowski, C. S. (1999). Introduction to the special issue: Work–family research in occupational health psychology. *Journal of Occupational Health Psychology*, 4(4):301.
- Wheatley, D. (2017). Employee satisfaction and use of flexible working arrangements. *Work, Employment and Society*, 31(4):567–585.
- Williams, J. C., Blair-Loy, M., and Berdahl, J. L. (2013). Cultural schemas, social class, and the flexibility stigma. *Journal of Social Issues*, 69:209–234.
- Zapf, I. and Weber, E. (2017). The role of employer, job and employee characteristics for flexible working time: An empirical analysis of overtime work and flexible working hours’ arrangements. Technical report, IAB-Discussion Paper.

A Additional results and figures (for online dissemination)

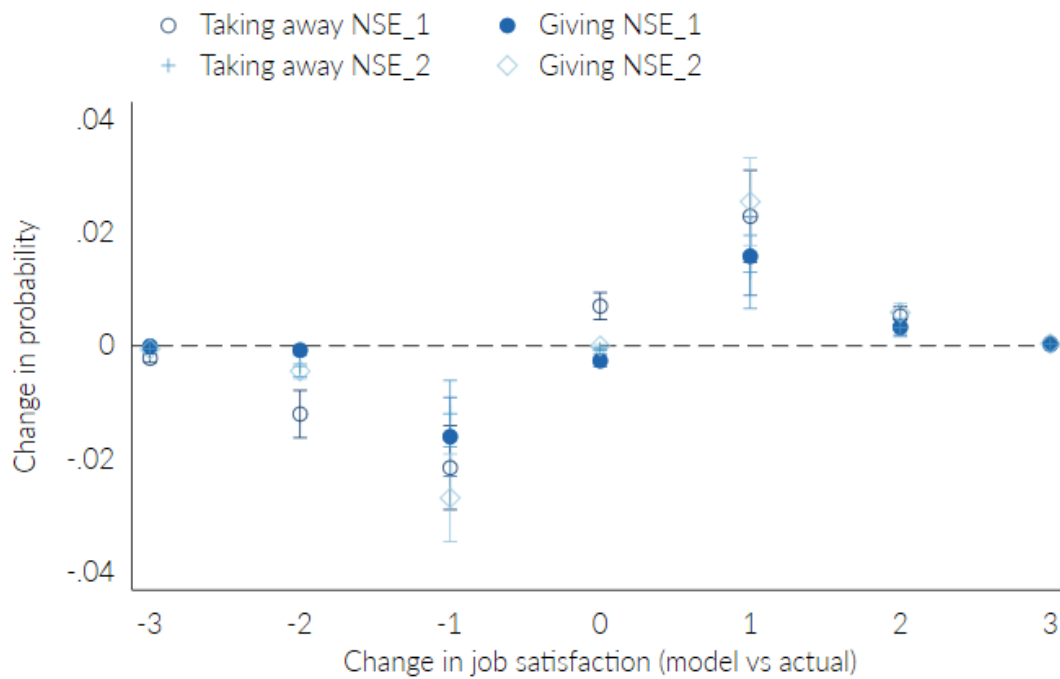
Figure A1: Job satisfaction in wave 6 of EWCS: comparing categorical and index measure



Notes The graph is based on EWCS data from waves 6. Job satisfaction index (JS_I) is factor-based measure utilizing five dedicated questions addressing various aspects of job satisfaction. Job satisfaction (JS) is a categorical self-reported variable.

Figure A2: Impact of gender on changes in job satisfaction: average marginal effects

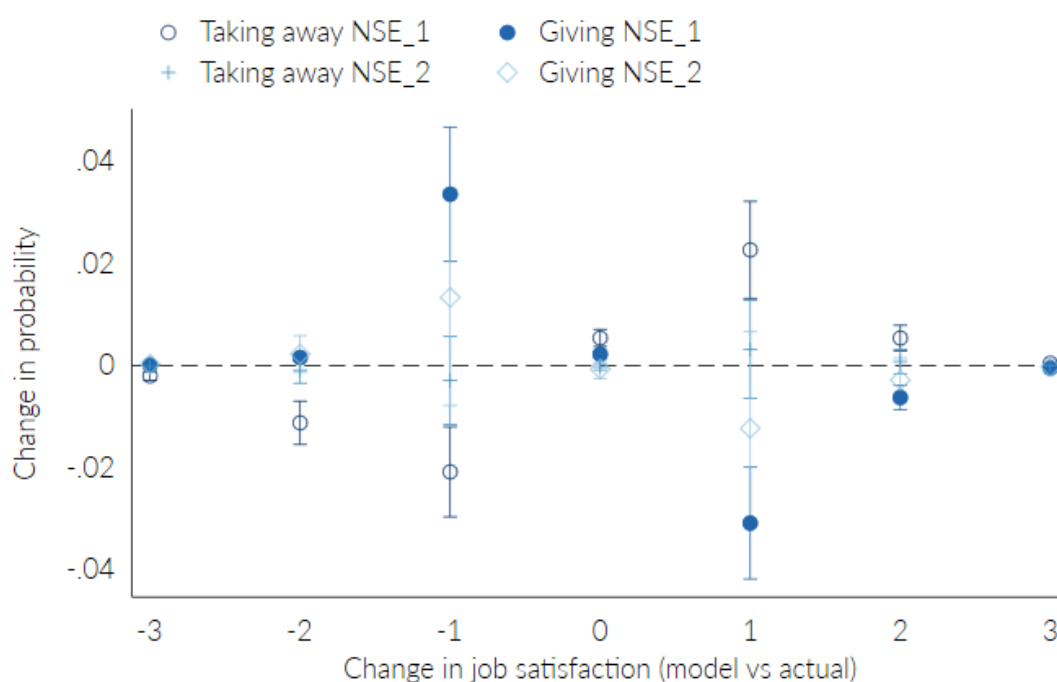
Marginal effects: women = 1



Notes Horizontal lines represent 95% confidence intervals. We observe a negative impact of being a woman in the case of a negative change in the job satisfaction, and positive impact in case of a positive change. That is, women are more likely than men to increase their job satisfaction when changing work flexibility regime. The result is common for all the models, i.e. it does not matter whether we assume irregular working arrangements for females with fixed schedules or vice versa, and consistent with the results in Table 4. The direct marginal effects interpretation is that women are equally likely to report extreme changes (-3 and +3) in job satisfaction index as men. However, women are around 2 percentage points (depending on the model) more likely to experience a positive change (of +1) in the job satisfaction after the change in work arrangements, and about 2 percentage points less likely to face a one level decrease in job satisfaction.

Figure A3: Impact of children on changes in job satisfaction: average marginal effects

Marginal effects: children under 7yo in hh =1



Notes Horizontal lines represent 95% confidence intervals. We only observe significant marginal effects for taking away NSE_1 and giving NSE_1 models. The direct marginal effects interpretation is that workers with children under 7 years old are equally likely to report extreme changes (-3 and $+3$) in job satisfaction index as workers without such children. However, workers with children are around 4 percentage points more likely to experience a negative change (of -1) in the job satisfaction, and about 4 percentage points less likely to have a positive change (of $+1$) by moving to irregular work arrangements (giving NSE_1). The outcome is consistent with the results in Table 4.