Understanding the Heterogeneity of Earnings Losses After Job Displacement: A Machine-Learning Approach*

Andreas Gulyas† Krzysztof Pytka‡

Current version:
April 2024
First version:
October 2019

Abstract

Applying a generalized random forest to Austrian administrative data, we uncover significant systematic heterogeneity in worker outcomes after job displacement. A quarter of workers face wage losses of 30%, while another quarter experience no losses or even gain. Among many factors, firm wage premia are the most important determinant of the wage loss heterogeneity, while workers' age is the most important for employment losses. Our findings suggest that earnings losses stem from mean reversion in firm wage premia rather than the destruction of firm-specific human capital. Compositional differences in individual characteristics account for most of the cyclicality in earnings losses.

Keywords: Job displacement, Earnings losses, Causal machine learning

*A previous version was circulated under the name of "Understanding the Sources of Earnings Losses After Job Displacement: A Machine-Learning Approach." Funding by the German Research Foundation (DFG) through CRC TR 224 (Project A3) and support by the state of Baden-Württemberg through bwHPC and the German Research Foundation (DFG) through grant INST 35/1134-1 FUGG is gratefully acknowledged. Johannes Gönsch, Michaela Paffenholz, and Tim Bartelsen provided excellent research assistance. We thank Clara Santamaría, Jan Stuhler, Stefan Wager and seminar participants at the National Bank of Austria, Carlos III, CRC TR 224 conference, NOeG Winter Workshop, Causal Machine Learning Workshop St. Gallen, SOLE/EALE, IZA, Oslo Macro Conference, CERGE-EI, Search and Matching Conference, and CRC Labor Market Workshop for helpful comments. All remaining errors are ours.

†University of Mannheim; `andreas.gulyas@uni-mannheim.de`

‡University of Mannheim; `pytka@uni-mannheim.de`

I. Introduction

A sizable literature documents that, on average, workers displaced during a mass layoff experience significant losses in annual income lasting over 15 to 20 years.[1] Motivated by this evidence, many labor market policies are designed to avoid job losses in general. Prominent examples include firm bailouts, employment protection, and employment subsidies such as short time work schemes, which are often applied in a non-discriminatory fashion. Critics argue that not all jobs are worth saving, and such programs should be, if anything, targeted. To design effective policy responses, we not only need to identify the consequences of job loss for the average worker, but also understand how earnings losses differ across individuals, and which pre-displacement characteristics lead to higher losses. This paper is the first to document how the long-term effects of job losses differ across all displaced workers in the population and identifies the main drivers of those losses.

To this end, we draw on recent advancements in machine learning to estimate heterogeneous treatment effects. To document the whole distribution of earnings, employment, and wage losses across different workers, we implement a generalized random forest by Athey *et al.* (2019) to a difference-in-difference (hereafter DiD) setup. Furthermore, we study how post-displacement employment histories differ across high- and low-loss individuals and decompose wage losses into a match component and losses in firm pay premia. Next, we identify which pre-displacement characteristics are the most important in determining the severity of the job losses' scarring effect. This way the machine learning procedure helps to distinguish between many competing theories about earnings losses in the literature.

Our approach estimates the causal cost of job loss nonparametrically as a function of observables. This is achieved by exploring possible data splits and choosing those ones for which between-group differences in losses are the highest. By recursively splitting the dataset into smaller and smaller subsamples, the algorithm builds a tree which detects heterogeneity in losses. To avoid detection of artificial heterogeneity, instead of growing a single tree, we build a random forest consisting of many tree-based models. Every tree is trained using a random subset of observations and variables and then their estimates are combined.[2] Another

[1] Jacobson *et al.* (1993), Neal (1995), Couch and Placzek (2010), Davis and Von Wachter (2011), Farber (2011), Farber (2017), among many others. Furthermore, job displacements have been shown to have detrimental effects on health (Schaller and Stevens, 2015), longevity (Sullivan and Von Wachter, 2009), and on children of displaced workers (Lindo, 2011; Rege *et al.*, 2011).

[2] This statement is for expositional reasons and might not be precise enough. Strictly speaking, for an individual characterized by observables $\mathbf{z}$, the random forest provides weights $\alpha_i(\mathbf{z})$ measuring similarity of all other observations indexed by i. Those weights are used in weighted linear regression to estimate the causal effect at $\mathbf{z}$. The whole algorithm is presented in Subsection IV.C in greater detail.

appealing property of random forests is that we are able to quantify standard errors arising from two sources, the machine-learning procedure and the estimation procedure.

We use the universe of Austrian social security records over three decades to examine the heterogeneity in earnings, employment, wage, and firm wage premia losses. Displaced Austrian workers suffer losses of similar magnitude compared to those displaced in Germany (Schmieder *et al.*, 2023) and the United States (Davis and Von Wachter, 2011). Specifically, ten years after job loss, displaced workers are 20 days less employed per year and persistently earn 8 log points lower wages than the control group.

Our algorithm documents losses conditional on workers' and job characteristics. For this reason, we construct 19 variables and feed them into our learning procedure. We seek to understand how earnings losses vary with these channels and to identify which of these are the most important ones. These variables are derived from the most prominent theories of earnings losses. We consider channels that tie earnings losses to losses in job-specific and general human capital, particularly good matches, job security, and firm wage premia. Firm wage premia are identified through firm fixed effects in a Mincerian wage regression following Abowd *et al.* (1999), and the match effect as the residual from a similar regression. In addition, we study how losses vary with various local labor market conditions, socio-demographic factors, and the business cycle.

To study the heterogeneity in earnings, employment, log-wages, and firm wage premia losses we train separate random forests. We document that the consequences of job losses are far from uniform. Over the 11 years we follow workers after the job loss, earnings losses range from over 15,000€ per year to even income gains. The most striking differences are in wage changes. While 25% of the workers face either little wage losses or even gains, the quarter of workers with the highest losses experience yearly wage declines of close to 30 percent. How do employment histories after job loss differ across high- and low-earnings-loss individuals? Those individuals identified to face the highest losses suffer also the largest declines in employer wage premia and match effects. In addition, they are the group that has the lowest propensity to remain in their original industry, move to more unstable firms, and have more job changes subsequently. In contrast, the group of workers which go through the job loss unscarred, move to better paying and more stable firms, increase the match effect, and face fewer subsequent job changes. All of the groups show little convergence in any of the considered labor market outcomes. This suggests that the changes experienced just after the job separation are permanent for all groups and prevail in the long run.

What pre-displacement worker and job characteristics determine the severity of losses?

Our methodology allows us to jointly compare many theories at once. First, we use a standard measure of variable importance from the machine-learning literature, which essentially counts how often a variable is used in the construction of the random forest. The pre-displacement firms' wage premium is by far the most important factor determining the level of earnings, wage, and firm wage losses, with all other factors only playing a second order role. In contrast, it does not play an important role for employment losses, where workers' age is the most important factor. We arrive at the same conclusions by decomposing how much each factor contributes to the variance of the treatment effects.

These conclusions are also underscored by computing how losses change by varying one channel at a time, holding all other variables constant at their median. The impact of certain variables is often judged by sample splitting, which does not allow to keep other correlated job characteristics fixed, potentially leading to spurious relationships. Ceteris paribus, workers separating from the lowest quintile of the firm pay distribution experience modest wage gains from job displacement, whereas workers separating from the best paying decile of firms face wage losses of 16 log points. The differences in log-wage losses mimic losses in firm wage premia, which show a mean-reversion pattern. Workers with above median firm wage premia lose in terms of firm pay, whereas the other half gain. Thus, for workers employed in well-paying firms, changes in firm pay explain a large fraction of wage losses, whereas for workers employed in average paying firms it explains very little. Consequently, compositional differences in the firm wage premia of displaced workers across samples could be a potential explanation why Lachowska *et al.* (2020) find that declines in firm pay explain little of wage losses in the state of Washington during the financial crisis, whereas Schmieder *et al.* (2023) finds the opposite for Germany.

We also find steep slopes in losses with respect to person fixed effect and worker's age, but the two operate through different channels. On one hand, older workers face negligible wage changes but high employment losses. On the other hand, earnings losses for workers with high person fixed effect before the displacement originate from depressed wages. We do not find that the cost of job loss varies much with all other factors. This is surprising given that lost job-specific human capital is one of the most prominent theories for earnings losses. We show that once we control for confounding factors, losses do not vary much with job tenure. All in all, our findings provide evidence that earnings losses can be understood by mean reversion in firm wage premia, rather than by a destruction of firm-specific human capital, while earnings losses for older workers are mostly driven by employment losses.

Prior research has shown strong cyclicality in earnings losses.[3] In contrast, we find that earnings losses are not affected by the business cycle directly. To understand this discrepancy, we use the fact that our methodology enables us to estimate earnings losses at the individual level and decompose the cyclical variation into a pure recession effect and compositional differences due to the fact that different workers are displaced during a recession than during an expansion. During recessions, the composition of displaced workers shifts towards worker and job characteristics that are associated with higher losses, which explains over 90% of the cyclicality. This highlights the importance of the ability of our machine-learning approach to hold confounding factors constant.

The most common approaches in the existing empirical literature to study heterogeneity in treatment effects are either parametrically by interacting the treatment indicator with many covariates, or by using quantile regressions. We juxtapose the heterogeneity of earnings losses identified through the random forest with these two methods. While the DiD regression with many interaction terms is successful in detecting heterogeneity, a large portion of it is caused by the estimation noise. Only 35% of observations are significantly different from the median estimate in comparison to 67% in the random forest. The quantile DiD regression detects significantly less heterogeneity than our random forest, which at first glance might look very surprising. The reason is that this class of methods is well suited for studying overall changes in the distribution of the variable of the interest. Nonetheless, they do not capture individual treatment effects very well, especially if ranks of the dependent variable before and after the treatment are not the same. This is the case in our application, as workers with high earnings before displacement typically face the highest earnings loses.

In addition to previously mentioned papers, we also contribute to the small but growing number of papers using machine learning to study heterogeneous treatment effects in economics. Examples include Davis and Heller (2020), who study the heterogeneous effects of youth employment programs, and Knaus *et al.* (2022) use a LASSO model to study treatment heterogeneity of job search programs. Gregory *et al.* (2021) use a k-means algorithm to classify workers into three clusters based on labor market histories and document the heterogeneity in earnings losses for these three groups. We document the whole distribution of losses and identify pre-displacement worker and job characteristics that are relevant for shaping earnings losses. In a recent working paper, Athey *et al.* (2023) use a similar machine-learning algorithm to ours to study heterogeneity in the cost of job loss in Sweden. In contrast to us, they find a limited impact of firm wage premia, and workers' age as the

[3] *E.g.* Davis and Von Wachter (2011), Schmieder *et al.* (2023).

most important variable accounting for the heterogeneity. Although speculative at the moment, one difference could be in the differential importance of employment and wage losses. In Austria, as in found in other settings (e.g. Schmieder *et al.* (2023)), long-term losses are mostly driven by wage losses, and not employment losses. In contrast, in Sweden long-term losses are mostly driven by employment losses, for which workers' age also plays the most important role in Austria. This highlights the importance of exploring the determinants of the heterogeneity of employment and wage losses together.

The rest of the paper is organized as follows. The next section describes the empirical setting in Austria, as well as the sample selection. Section III presents the average cost of job displacement. Section IV describes the machine-learning algorithm used to identify the driving forces behind earnings losses. Section V documents heterogeneous scarring effects of job displacement and section VI discusses the sources behind earnings losses. The last section concludes.

II. Empirical Setting

We use the administrative employment and unemployment records from the social security administration in Austria from 1984 through 2019. This data comprises day-to-day information on all jobs and unemployment spells covered by social security in Austria (Zweimüller *et al.*, 2009). Thus, it contains all private sector jobs and a large fraction of public sector employment, but excludes self-employed and public servants who are not covered by social security. It contains information on yearly earnings for each worker-establishment pair, and basic socio-demographic information at the worker level such as age, gender, a flag for a blue collar occupation, and citizenship.[4] Each establishment (we use firm and establishment exchangeably from here on) has a unique identifier, we have information on its geographic location, age, and 4-digit industry classifier.

A. Definition of Job Displacement and Mass Layoff

To ensure comparability with the previous literature on displaced workers, we follow the typically applied definitions and sample restrictions as much as possible. Workers are considered displaced if they separated from their primary employer that experienced a mass layoff in the given year. We define a mass-layoff event at the firm level in year t if the size of the firm declined by more then 30 percent during year t. To avoid selecting volatile firms,

[4]We deflate all earnings to 2017 level using the CPI index provided by the Austrian Statistical Agency.

we exclude firms that grew by more than 30 percent in either $t-1$, or $t-2$, as well as firms that are larger 3 years after the event than before. To have a meaningful measure of firm growth, we only consider establishment with at least 30 employees. To avoid misspecifying mergers, outsourcing, or firm restructures as mass-layoffs, we compute a worker cross flow matrix for all firms in each year. We exclude all firms where more than 20 percent of their workforce ends up working for the same employer in $t+1$.[5] Thereby we exclude mass layoff firms with large worker flows to other firms. Not correcting for these potential measurement errors might lead to a significant underestimation of earnings losses. We also disregard mass layoff events from the public administration (Nace, Level 1, Code O).

B. Sample Construction

We depart from most of the earnings loss literature and do *not* restrict our sample to males only. Specifically, we are interested in how earnings losses vary across different individuals based on their characteristics, including gender. We proceed by selecting everybody who is employed on the reference day of January 1st each year. This results in 50,708,644 person-year observations. We follow the literature and restrict our sample to workers aged 24-50, employed at a firm larger than 30 employees, and with job tenure longer than 2 years on the reference day. We remain with 13,311,284 person-year observations. As is apparent from the reduced sample size, these two sample selection criteria are restrictive, but at the same time necessary to cleanly identify an unexpected, no-fault job loss. A common critique is that these sample restrictions are selecting workers that are bound to face high earnings losses. We will use our machine-learning algorithm to make out-of-sample predictions on earnings losses for the general population.[6]

Out of these remaining observations, we define a person to be displaced in t if a worker separates from a firm experiencing a mass layoff, and the worker is not reemployed at the same firm at any point in the next 10 years. If a worker suffers multiple mass layoffs, we only consider the first one. We identify 59,144 displaced worker events between 1989 and 2009.

Some workers disappear from our dataset over time. This happens on the one hand because workers might not find employment subject to social security insurance anymore and drop out of the labor force. On the other hand, this could also happen if workers move

[5] For an in-depth discussion, see Hethey-Maier and Schmieder (2013).

[6] Surprisingly, the distribution of losses for mass-layoff workers and the general population is comparable, as discussed in more detail later. The reason behind this is that the variables of the sample selection turn out not to be first-order determinants of earnings losses.

	Displaced	Selected Control Group	Not Selected
Age	37.72	37.72	37.88
$\log(w_{t-1})$	4.47	4.47	4.58
$\log(w_{t-2})$	4.45	4.45	4.55
Job Tenure (in days)	2381.19	2373.98	2587.40
Manufacturing	0.55	0.55	0.45
Firm Size	212.92	206.03	814.50
Female	0.40	0.41	0.38
Obs	59,144	59,144	13,183,167

Table 1: Sample characteristics of displaced workers compared to the selected control group via propensity score matching and the universe of worker/year observations satisfying sample restrictions

into self-employment or move abroad. We decided to use only information on workers who either have an employment spell covered by social security or a registered unemployment spell in a given year. This likely underestimates the true costs of job-displacement, as we do not measure losses associated with dropping out of the labor force.

Note that we do not restrict the control group to have stayed at their employer after t. The potential comparison group consists of non-displaced workers subject to the same sample restrictions. This includes workers employed at firms without any mass layoff event during year t or workers in mass layoff firms who did not separate.

Non-displaced workers may differ in many characteristics from the displaced workers, as can be seen in Table 1. Following many papers in the literature, e.g. Schmieder *et al.* (2023); Bertheau *et al.* (2023); Halla *et al.* (2018), we use propensity score matching in order to obtain a control group that is as similar as possible to displaced workers. In each year, for all workers satisfying the sample restrictions, we estimate the propensity to experience a displacement event as a function of the following worker and firm characteristics: worker's log-wage in year $t-1$ and $t-2$, tenure, age, establishment size in year t, as well as a dummy for working in the production sector.[7] For each displaced worker in a given year, we select the non-displaced worker with the nearest propensity score without replacement. Table 1 shows that our matched control group is very similar to displaced workers in observable characteristics. The two groups are virtually indistinguishable in terms of pre-displacement evolution of earnings, days employed, and log-wages, as can be seen in Figure 15 in the appendix, which plots raw averages across these groups over time.

[7] We also experimented with different sets of matching variables, all of which lead to similar results.

III. The Average Cost of Job Displacement

Throughout the paper, we are interested in identifying how the cost of job loss differs across individuals. We nevertheless start by discussing the estimation strategy for the homogeneous treatment case, which will be extended to heterogeneous treatment effects next section. The average cost of job loss can be compactly estimated using the following DiD setup:

$$y_{it} = \tau \mathbb{1}(t \geq t^*) \times D_i + \theta D_i + \gamma_t + \epsilon_{it}, \tag{1}$$

where D_i is an indicator equal to one for a displaced persons, t^* the displacement year and t the current year. The period fixed effects γ_t measure the evolution of the left hand side variable of the control group, and D_i absorbs initial differences in labor market outcomes between the control and treatment group. τ measures the average yearly cost of job loss in the 11 years following job displacement, relative to the counterfactual of no job loss. Table 2 shows the estimates from this regression for different specifications. Column (1) reports the estimates for equation (1) without any controls, in column (2) and (3) a polynomial in age and worker fixed effects are added. In columns (4) to (6) we additionally report the cost of job loss estimates from the following event study equation

$$y_{it} = \sum_{j=-4}^{10} \delta_j \mathbb{1}(t = t^* + j) \times D_i + \theta D_i + \gamma_t + \epsilon_{it}, \tag{2}$$

where the average cost of job loss is given by the average of all post-displacement coefficients δ_j, i.e. $\tau = \frac{1}{11}\sum_{j=0}^{10} \delta_j$.

In all specifications, the yearly earnings losses amount to close to €6,500 per year. Because of the computational burden of the machine learning algorithm, we will not able to include worker fixed effects as controls as is often done in the literature. But the results from Table 2 show that because our matching procedure selected very similar workers as a control group, adding worker fixed effects or a polynomial in age does not significantly change the estimated costs of job displacement.

	Dependent variable:					
	Yearly Income					
	(1)	(2)	(3)	(4)	(5)	(6)
τ	-6444.5	-6653.1	-6656.5	-6360.0	-6732.42	-6725.67
	(65.3)	(63.6)	(61.0)	(81.1)	(79.01)	(74.58)
DiD	✓	✓	✓			
Event Study				✓	✓	✓
Worker FE		✓	✓		✓	✓
f(age)			✓			✓
Observations	1746289	1746289	1746289	1746289	1746289	1746289
R^2	0.04	0.7	0.7	0.04	0.7	0.7
Adjusted R^2	0.04	0.7	0.7	0.04	0.7	0.7

Table 2: Estimates of earnings losses with different sets of controls. DiD regressions based on equation (1) and Event Study regressions based on equation (2). The event study coefficient is given by the averages of the post-treatment coefficients, i.e. $\frac{1}{11}\sum_{j=0}^{10}\delta_j$.

IV. Heterogeneity in the Cost of Job Loss – Machine Learning Approach

The goal of our exercise is to identify the heterogeneity in earnings losses and its sources. To this end, we employ a machine learning procedure built on the methodology of generalized random forests, recently developed by Athey *et al.* (2019). The average cumulative earnings losses after job displacement can be estimated directly from formula (1) using standard statistical methods. Nonetheless, this approach has some serious limitations. For instance, this type of the estimate provides the average treatment effect and does not provide information on its underlying heterogeneity. If the earnings losses are not uniform across individuals, then the individual scarring effects may be a far cry from the estimated average.

Conceptually, our approach consists in estimating the version of equation (1) with heterogeneous scarring effects:

$$y_{it} = \tau(\mathbf{z}_i)\mathbb{1}(t \geq t^*) \times D_i + \theta(\mathbf{z}_i)D_i + \gamma_t(\mathbf{z}_i) + \epsilon_{it}, \tag{3}$$

where $\mathbf{z}_i$ are the values of observable variables (henceforth called *partitioning variables*) for individual i. $\tau(\mathbf{z}_i)$ is the function that describes how the cost of job loss changes with individual worker and job characteristics $\mathbf{z}_i$. The functional specification of $\tau(\mathbf{z}_i)$ is assumed to be unknown. To uncover the true form of $\tau(\mathbf{z}_i)$, we employ a generalized random forest by Athey *et al.* (2019), adapted to a DiD setting.

We proceed by describing the outcome variables and partitioning variables, which are associated with the most prominent channels from the job displacement literature. Details on how we implement generalized random forests for our application are described subsequently in Subsection IV.C. There are several important advantages of our approach in comparison to traditional techniques. In Section VII, the findings from our random forest are juxtaposed with the most popular methods used in the empirical literature.

A. Outcome Variables

We study the cost of job loss in terms of four dimensions. First, we estimate how yearly earnings in Euro change after job displacement. Second, we measure employment losses as the change in days employed covered by social security. Third, we study how worker's log daily wage at their dominant employer evolves after job loss conditional on re-employment.[8] Fourth, we analyze to what extent the decline in log-wages can be explained by transitions to lower paying firms. We estimate firm wage premia following Abowd *et al.* (1999), which has become the workhorse model for empirically estimating the firm pay component. We estimate

$$\ln(w_{it}) = \psi_{J(i,t)} + \alpha_i + \theta_t + x_{it}\beta + \epsilon_{it}, \tag{4}$$

where $\ln(w_{it})$ is the log daily wage of the dominant employer at period t, $\psi_{J(i,t)}$ represents the establishment fixed effect of the employer of worker i at period t, α_i the worker fixed effect, θ_t the year fixed effect, and x_{it} are time varying observables, comprising of a cubic polynomial of age.

In order to have one estimate per firm which is comparable across time, we use the universe of private sector employment spells in Austria from 1984 through 2019 in the estimation. In order to avoid endogeneity issues, we disregard individuals in our earnings loss sample from the computation of firm fixed effects.[9] The advantage of this approach is that losses in firm wage premia are measured in log-wages, and thus directly comparable to the

[8]We compute daily wages by dividing yearly income from the employer with the highest earnings in a given year and divide it by the number of days employment at this employer.

[9]Firm fixed effects are identified from wage changes of workers moving across firms. Thus, if workers experience earnings losses in mass-layoffs, the mass-layoff firm will be estimated to be a high fixed-effect firm

wage losses.[10] The details for the estimation are found in Appendix C.

B. Partitioning Variables

What are the reasons that earnings losses differ across individuals? On the one hand, workers in different environments face different re-employment probabilities after job loss, on the other hand, displaced workers might face different wage declines after re-employment. The workhorse framework in empirical research typically posits that wages consists of multiple components: $ln(w_{it}) = \alpha_i + \psi_{J(i,t)} + \beta * tenure_{it} + \varepsilon_{it}$. α_i captures the compensation a worker receives for their skills, $\psi_{J(i,t)}$ the firm specific wage component, $\beta * jobtenure_{it}$ the enumeration for the worker's job specific human capital accumulated over their job tenure, and ε_{it} represents an idiosyncratic match component.

In job ladder search models (McCall, 1970; Cahuc *et al.*, 2006), laid off workers fall down the job ladder in terms of firm wages and match quality and have to climb back through job-to-job transitions. Thus, earnings losses can vary first, because workers with higher firm wage components or better previous match quality face a steeper fall off the job ladder, and hence higher immediate wage losses. In addition, workers with high skills α_i before displacement have more room to face depreciation of these skills and might be more susceptible to wage declines. Moreover, workers with higher fixed effects might have accumulated more savings and higher claims to unemployment benefits, which would enable them to be pickier in their job search. This would lead to higher employment losses, but lower wage losses after job loss (Nekoei and Weber, 2017). We study all these channels by estimating equation (4) on a rolling window of all social security spells in the five years before the event, and then constructing deciles of the estimated firm fixed effect $\hat{\psi}_{J(i,t)}$ and worker fixed effect $\hat{\alpha}_i$, and the wage residual $\hat{\varepsilon}_{it}$. This way we strictly use pre-displacement observations to estimate these channels.

Perhaps the most dominant theory of earnings losses is that they reflect losses in firm-specific human capital. Workers with higher prior tenure accumulated more of this capital and are therefore more susceptible to high earnings losses. This channel is captured by the inclusion of job tenure as a partitioning variable.

All these factors might be compounded if workers also lose particularly stable jobs (Jarosch, 2023). Earnings losses are then higher, as the comparison group is less likely to lose their job and more likely to hold on to the favorable job characteristics and are less

[10] Table 6 in Appendix C shows that the variation in the firm fixed effect alone explains around 30% of the variation of the log-wages.

likely to face human capital depreciation during unemployment. We consider this channel by adding the average firm level separation rate in the five years leading up to the mass layoff event as a partitioning variable.

In addition, earnings losses might vary because workers face different job finding probabilities or job offer distributions. The latter is captured by computing the average percentile of the firm wage effect in the local labor market.[11] We include a number of factors that potentially affect job finding probabilities: In addition to the average unemployment rate over our study period in the local labor market and in the prior industry,[12] we study how the business cycle affects earnings losses by constructing a recession dummy indicating years with high unemployment.[13] Workers in concentrated labor markets or separating from a monopsonist employer might face more difficulty to quickly find a similarly good job. Thus, we include the overall local labor market concentration as measured by the Herfindahl-Hirschman index, the firm size, and the employment share of the previous employer as partitioning variables.

Workers with higher job mobility in the past might find it easier to climb back the job ladder. The number of previous held jobs, which proxies for past job mobility is therefore included as an additional partitioning variable. Furthermore, we include a number of socio-demographic factors such as worker age, a dummy for Austrian citizenship, gender, in addition to blue collar occupation, working in the manufacturing sector, and firm's age, as the job finding probabilities and thus earnings losses might vary with these factors as well.

To enhance interpretability, we categorize all continuous variables into deciles according to the overall distribution of Austrian workers on our reference day, and not only the selected displaced worker sample. This way the heterogeneity is easily interpretable in terms of the overall employment distribution in Austria. Table 7 in the appendix summarizes all the definitions of the partitioning variables, Table 8 shows the cutpoints for the deciles, and Figure 16 reports a correlogram of all partitioning variables. Earnings losses are likely a combination of all these factors, and different channels interact with each other. In the next section we describe the applied machine learning algorithm that enables us to disentangle the contribution of all these different channels.

---

[11] Specifically, we compute the average firm wage premia of all jobs in a given region leaving out all jobs of the worker's current employer. Formally for every worker i employed at firm $J(i,t)$ we compute $\sum_{k \notin J(i,t) \wedge k \in r(i)} \hat{\psi}_{J(k,t)} / \#(k \notin J(i,t) \wedge k \in r(i))$, where $r(i)$ is the region of the worker i.

[12] We use the NUTS-3 district (35 categories) for the local labor market and NACE level-1 industry classification (21 categories) of employers.

[13] We define a mass layoff occurring in a recession if the unemployment rate is above it's trend in the year of the mass layoff event. Because the unemployment rate is trending upward in Austria since three decades, the deviation from its trend is a better measure of the business cycle than the level of unemployment.

C. Bird's-Eye View of Machine Learning Algorithm

We now describe the implemented method of a generalized random forest (Athey *et al.*, 2019). We seek to estimate the cost of job loss locally at worker and job characteristics $\mathbf{z}$ from equation (3), which is characterized by following local moment conditions:

$$\mathbb{E}(\mathbf{x}'_{it}\varepsilon_{it}|\mathbf{z}) = \mathbf{0}_{18}, \tag{5}$$

where $\mathbf{x}'_{it} = [\mathbb{1}_{(t\geq t^*)}D_i, D_i, \mathbb{1}_{\{t=-5\}}, \cdots, \mathbb{1}_{\{t=0\}}, \cdots, \mathbb{1}_{\{t=10\}}]$, ε_{it} is the error term from (1), and $\mathbf{0}_{18}$ is a row vector with zeros of length 18. Our approach consists of defining similarity weights $\alpha_{it}(\mathbf{z})$, which measure the relevance of observation it to estimating the cost of job loss at $\mathbf{z}$, and estimating equation

$$\begin{aligned}(\tau(\mathbf{z}), \theta(\mathbf{z}), \boldsymbol{\gamma}(\mathbf{z})) &= \operatorname{argmin}_{\{\tau,\theta,\gamma\}} \left(\frac{1}{NT}\sum_{i}^{N}\sum_{t}^{T}\alpha_{it}(\mathbf{z})\mathbf{x}'_{it}u_{it}\right)\left(\frac{1}{NT}\sum_{i}^{N}\sum_{t}^{T}\alpha_{it}(\mathbf{z})\mathbf{x}'_{it}u_{it}\right)', \\ &\quad \text{s.t.}\forall_{i,t}u_{it} = y_{it} - \tau\mathbb{1}(t\geq t^*)\times D_i - \theta D_i - \gamma_t. \end{aligned} \tag{6}$$

Notice that the problem (6) takes weights $\alpha_{it}(\mathbf{z})$ as given and is solved for each value of partitioning variables $\mathbf{z}$ separately. For constructing these weights, a generalized random forest is used. For exposition purposes, the algorithm is presented in three steps. First, we show how a single tree in the spirit of Breiman *et al.* (1984) with a modified splitting criterion borrowed from Athey *et al.* (2019) is grown. Next, the approach is augmented to generalized random forests. Finally, we present the detailed numerical implementation and how the weights can be recovered from our random forest.

C.1. Tree Construction

The tree-based procedure consists in partitioning the dataset into smaller subsamples in which individuals exhibit similar earnings losses and at the same time the differences in earnings losses between subsamples are maximized. The data fragmentation is carried out using a sequence of complementary restrictions on partitioning variables. Due to the computational complexity, a top-down, greedy approach is traditionally used. The procedure of building a tree can be characterized in the recursive way by Algorithm (1). In each data partition (called also a node or a leaf) the scarring effect is estimated from equation (1) separately.[14]

Algorithm 1 Tree Algorithm of Recursive Partitioning

i. Start with the whole dataset and consider it as one large data partition, $\mathcal{P}$.

ii. For each partitioning variable z_k and its every occurring value $\overline{z}$, split partition $\mathcal{P}$ into two complementary sets of individuals i such that $\mathcal{P}_l = \{i \in \mathcal{P} : z_{ki} \leq \overline{z}\}$ and $\mathcal{P}_r = \mathcal{P} \setminus \mathcal{P}_l$ and estimate cumulative earnings losses τ_l and τ_r for both partitions by running two separate regressions of form (1) on $\mathcal{P}_l$ and $\mathcal{P}_r$.

iii. Choose the variable z_k and value $\overline{z}$ that maximizes:

$$(\tau_l - \tau_r)^2 \, \frac{n_l \cdot n_r}{N^2}, \tag{7}$$

where n_l and n_r are sizes of $\mathcal{P}_l$ and $\mathcal{P}_r$ and N is the sample size of $\mathcal{P}$.

iv. If (7) is smaller than a tolerance improvement threshold, then stop. Otherwise, go to step (ii) and repeat the splitting procedure for $\mathcal{P}_l$ and $\mathcal{P}_r$ separately, where $\mathcal{P}_l$ and $\mathcal{P}_r$ are new partitions subject to the splitting procedure, $\mathcal{P}$.

The main difference of our procedure from the textbook CART algorithm Breiman *et al.* (1984), is the splitting criterion. In the original approach, the algorithm aims at building a tree minimizing the squared sum of residuals.[15] In our setup, we are interested in growing a tree that explores the underlying heterogeneity of earnings losses between partitions of individuals with different characteristics. For this reason, we adapt the criterion (7) proposed by Athey *et al.* (2019). This criterion maximizes between-group differences of earnings losses, $(\tau_l - \tau_r)^2$, with an adjustment for more balanced splits, $\frac{n_l \cdot n_r}{N^2}$.

In applied economics, one alternative to splitting the dataset is to assume that the data-generating process is known and given and to estimate according to that process. In our application this would mean that we make an arbitrary decision upon the specification of $\tau(\mathbf{z}_i)$. However, in our strategy we are upfront about our agnosticism on $\tau(\mathbf{z}_i)$ and employ Algorithm (1) to learn the true specification. As a result, the learning procedure does three things: (i) chooses which variables are important and contribute to accounting for the heterogeneous scarring effects and which do not; (ii) detects non-linear relationships between τ and $\mathbf{z}_i$; (iii) detects interactions (including interactions of higher orders) between partitioning variables.

Figure 1 depicts a tree grown using the described algorithm. Every node is labelled with

[14] It is noteworthy that while all parameters from (1) are estimated, only the parameter of our interest, the scarring effect, is used in the splitting criterion (7).

[15] In regression trees, the squared sum of residuals is defined as $\sum_j \sum_{i \in \mathcal{D}_j} (y_i - \overline{y}_{\mathcal{D}_j})^2$, where $\mathcal{D}_j$ is a data subset obtained through sample partitioning procedure and $\overline{y}_{\mathcal{D}_j} = \frac{1}{|\mathcal{D}_j|} \sum_{i \in \mathcal{D}_j} y_i$ is the mean of the response variable in the specific set of data.

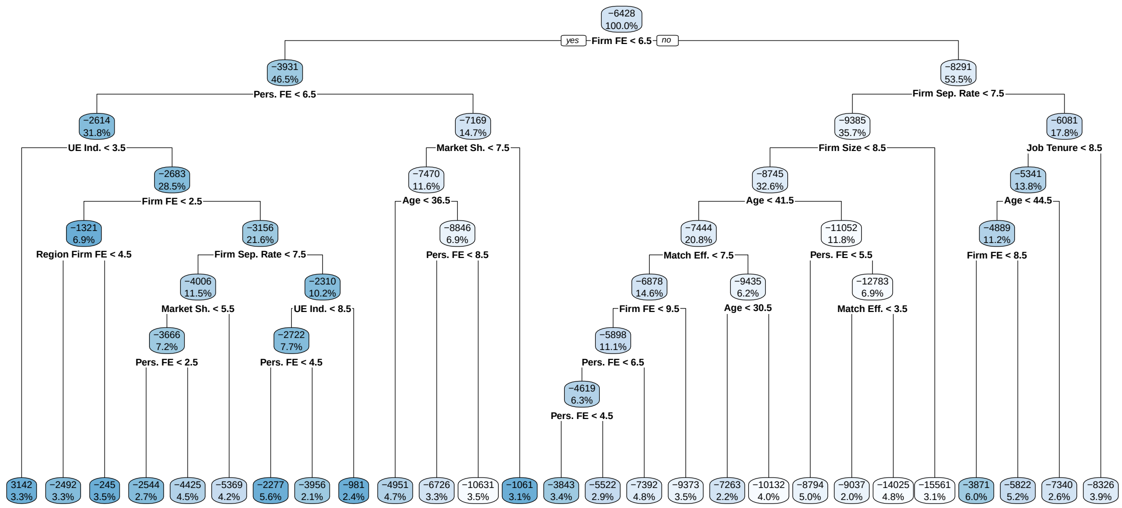


Figure 1: Heterogeneous treatment effect of job displacement. Tree was built with a CART algorithm using the relabeling strategy as described in the text. Minimum node size of 35,000 person-year observations.

the average earnings losses and the overall fraction of observations in the node. On the top, there is the root node containing all observations, and on the bottom there are final nodes that are not subject to further partitioning. Fractions of all final nodes (leaves) sum to 100%. In the whole dataset the average earnings loss is equal to around €6400. In the first iteration, the split that maximizes heterogeneity between groups goes according to the firm fixed effect. Individuals displaced from firms paying above the median experience losses of almost €8,300, while the workers separating from firms paying below the median suffer losses lower than 4,000. Overall, the partitioning procedure generates 27 final nodes endogenously. The number of binary splitting conditions used to generate each leaf varies from 3 to 8. Already at this point, we can observe that individuals displaced from high fixed-effect firms face higher losses than others. In addition, those losses can be amplified even more if, for instance, displaced workers had relatively stable jobs or exhibited a person fixed effect above the median level.

C.2. Generalization to Random Forests

One important advantage of tree-based models is its easy and very intuitive graphical illustration. Unfortunately, it is well-known that estimates can be non-robust and it is difficult to properly estimate their standard errors. Random forests proposed by Breiman (2001) are a refinement of the baseline method that address the typical concerns of tree-based models. The general idea behind random forests is quite easy and relies on building many trees through bootstrapping data observations. Moreover, in each split decision a subsample of considered variables is drawn. Consequently, the ensemble of decorrelated trees is grown, which means that the trees differ from each other and are built with different variables. This way only relationships that consistently show up in different bootstrapped samples are identified. In addition to this, each tree has been built using a so-called "honest" approach. This means that half of the bootstrapped sample was used to determine conditions which constitute data partitions, while with the other half the scarring effects were estimated in those partitions.[16] With the forest at hand, we can proceed with the construction of weights. Intuitively, the weights measure how often individuals fall into the same partitions across all trees. Then, those weights are used to solve (6) for each value of partitioning variables $\mathbf{z}$ separately. Formal derivation of the weights and how standard errors for individual treatment effects are computed using bag of little bootstraps as in Sexton and Laake (2009) are relegated to Section D of the Appendix.

V. Heterogeneous Scarring Effects of Job Loss

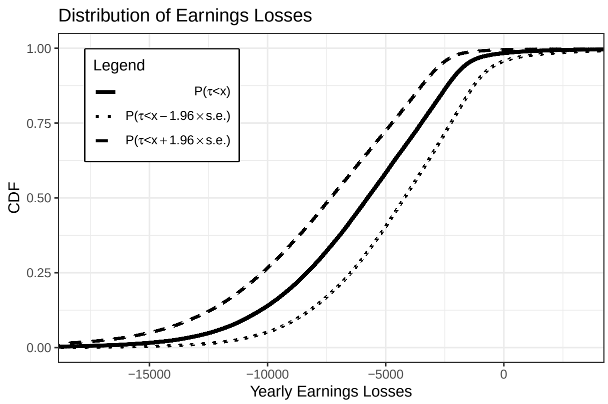
In this section we use the machine learning algorithm to document the unequal consequences of job loss across workers. We focus on four labor market aspects: losses in yearly income, losses in days employed, wage losses conditional on re-employment, and firm wage premia losses.

To analyze the losses in each of the four variables, we grow a forest consisting of 10,000 trees, with the minimum leaf size equal to 160 person-year observations. In each considered split, we draw 7 randomly chosen partitioning variables. Observations are clustered at the worker level.

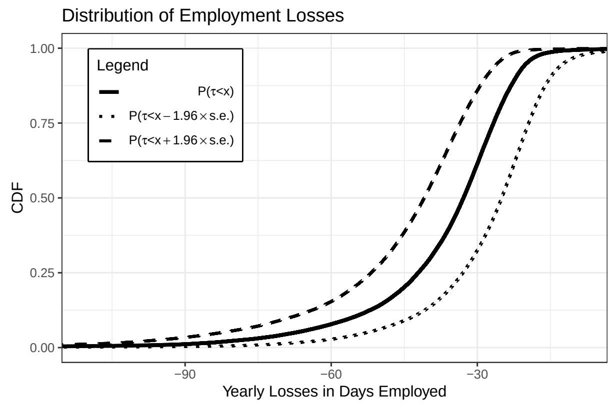
A. The Distribution in the Cost of Job Loss

The generalized random forest provides an estimated treatment effect for each individual worker based on their characteristics. We are thus in a position to document how the scarring effect of job displacement differs across workers. We start with plotting the distributions

[16] Thanks to this procedure we make sure we do not document spurious heterogeneity. If by any chance some splits are made due to some outliers, the estimated treatment effects are not affected by this.

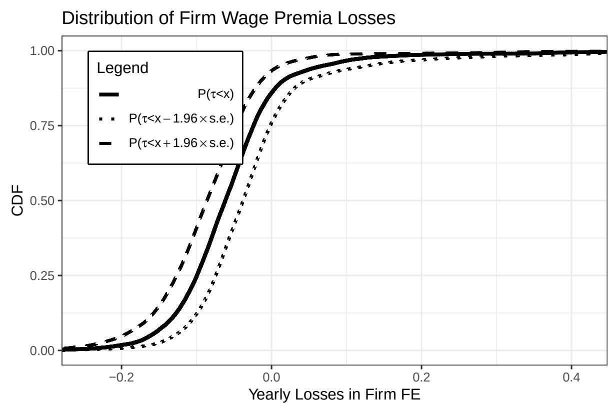


(a) Absolute Losses (in Euros)



(b) Employment Losses

(c) Losses in Log-Wages



(d) Losses in Firm Wage Premia

Figure 2: Cumulative Distribution Functions the losses in earnings (a) days employed (b), wage (c), and firm wage premia (d). Estimates from a generalized random forest. 0.05% outliers on both extremes not shown for better readability.

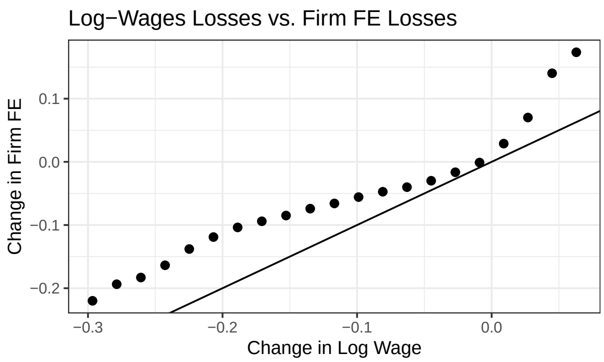
of τ_i for earnings, employment, wage, and firm wage premia losses. These treatment effects measure by how much the variable of interest changes for displaced workers relative to the control group on average over the 11 years following the displacement event. Figure 2 shows the cumulative distribution functions of the cost of job loss for the four variables. The figure clearly shows that the long-term consequences of job losses are far from uniform across individuals. While the median worker in our sample is facing income losses of €5,500 per year, distribution of individuals earnings losses ranges from losses of more than €17,500 per year to no losses at all. This heterogeneity is not driven by noisy estimates of the individual treatment effect. At the 95% confidence level, we estimate that 66% of workers face significantly different earnings losses compared to the median losses of €5,500.

Earnings losses are results of both forgone income during non-employment and declines in wages after reemployment. The latter itself potentially a result of losses in firm wage premia. In Panels (b), (c), and (d) of Figure 2, we plot the distributions of employment losses, log-wage losses, and firm wage premia losses. Again, there are striking differences in scarring effects across individuals. Unsurprisingly, nobody is estimated to gain days employed due to job displacement. Employment losses in the 11 years following mass layoff event are concentrated between 20 days and 50 days per year, 85 percent of workers face employment losses in this range. But there is a long left tail of very severe employment losses up to more than 90 days per year.

As seen in Panel (c), after re-employment, a quarter of workers experience sizable wage gains or losses of less than 5%, while a quarter of workers are facing wage declines by more than 15%. Firm wage losses upon job displacement have recently received a lot of attention. Lachowska *et al.* (2020) and Schmieder *et al.* (2023) study how much losses in firm wage premia explain wage losses of displaced workers and arrive at opposite conclusions. While a large faction of wage losses can be explained by losses in firm wage premia in Germany, they only play a small role in explaining wage losses in the state of Washington. Panel (d) shows that these two findings are not necessarily contradictory, because there are large differences across workers. In our sample a quarter of workers either gain in terms of firm pay, or face no significant losses. In contrast, 25% of workers lose more than 10% in terms of firm pay. Therefore, small compositional differences in the sample of workers can explain differences in losses of firm pay.

We study this in even more detail in Figure 3, where we present how wage losses at the individual level are related to the estimated losses in earnings, employment and firm wage premia. Panel (a) shows that there is a strong relationship between log-wage losses and losses in firm pay, suggesting that losses in firm pay explain a lot of variation in wage losses. Since both variables are measured in log-points, if losses in firm pay would fully explain wage losses, the relationship should perfectly line up along the 45 degree line, which is depicted as the solid line. In contrast, the figure shows that workers in general face higher wage losses than firm pay losses, the fraction of wage losses explained by firm wage premia losses ranges from 50 to 70 percent (conditional on wage declines). We will revisit this question in more detail in Section VI, where we will study how wage and firm wage losses depend on the previous employer's characteristics.

Panel (b) shows that higher wage losses are typically compounded by higher employment losses. Workers gaining in terms of wages are facing employment losses of less than 30 days



(a) Wage Losses vs Firm Pay Losses

(b) Wage Losses vs Employment Losses

(c) Wage Losses vs Earnings Losses

Figure 3: Bin scatter plot of the relationship between wage losses, and losses in firm wage premia, employment losses, and earnings losses. Estimates from generalized random forests.

per year, whereas workers with wage losses above 20 percent experience almost twice as high employment losses. In Section E of the Appendix, we will more formally decompose how employment and wage losses determine overall income losses.

Finally, Panel (c) plots the relationship between earnings losses measured in Euros and log-wage losses. It is clearly visible, that workers with higher absolute losses in earnings are losing more in wages in relative terms, with the correlation being 88% between the two. Thus, the heterogeneity of earnings losses is not just a results of heterogeneity in earnings before job displacement.

A long standing concern in prior research is the representativeness of the results to the general population. To tackle this, we use our grown forest to predict earnings losses for each employed worker given their worker and job characteristics.[17] Section G in the Appendix shows that the distribution of predicted losses for the general population is comparable to

[17] For workers who have lower job tenure or are employed at smaller firms than our sample restrictions, we assume the lowest possible values.

the distribution of losses among displaced workers.

Measuring the accuracy of the losses identified through the random forest is not an easy task. The reason for this is that we are not interested in accurate predictions of an observed dependent variable on the left hand side in Equation 3, but rather precise identification of an unknown treatment effect, one of the estimated regression coefficients on the right hand side. We address this concern by proposing an alternative evaluation procedure. In our exercise we create 50 distinctive groups.[18] For each group we estimate the average cost of job loss using Equation 1. Then we compare these treatment effects with forest-implied treatment effects. As we show, both measures are very highly correlated, which suggests that our method provides accurate estimates. We relegate more detailed discussion to Section F in the Appendix.

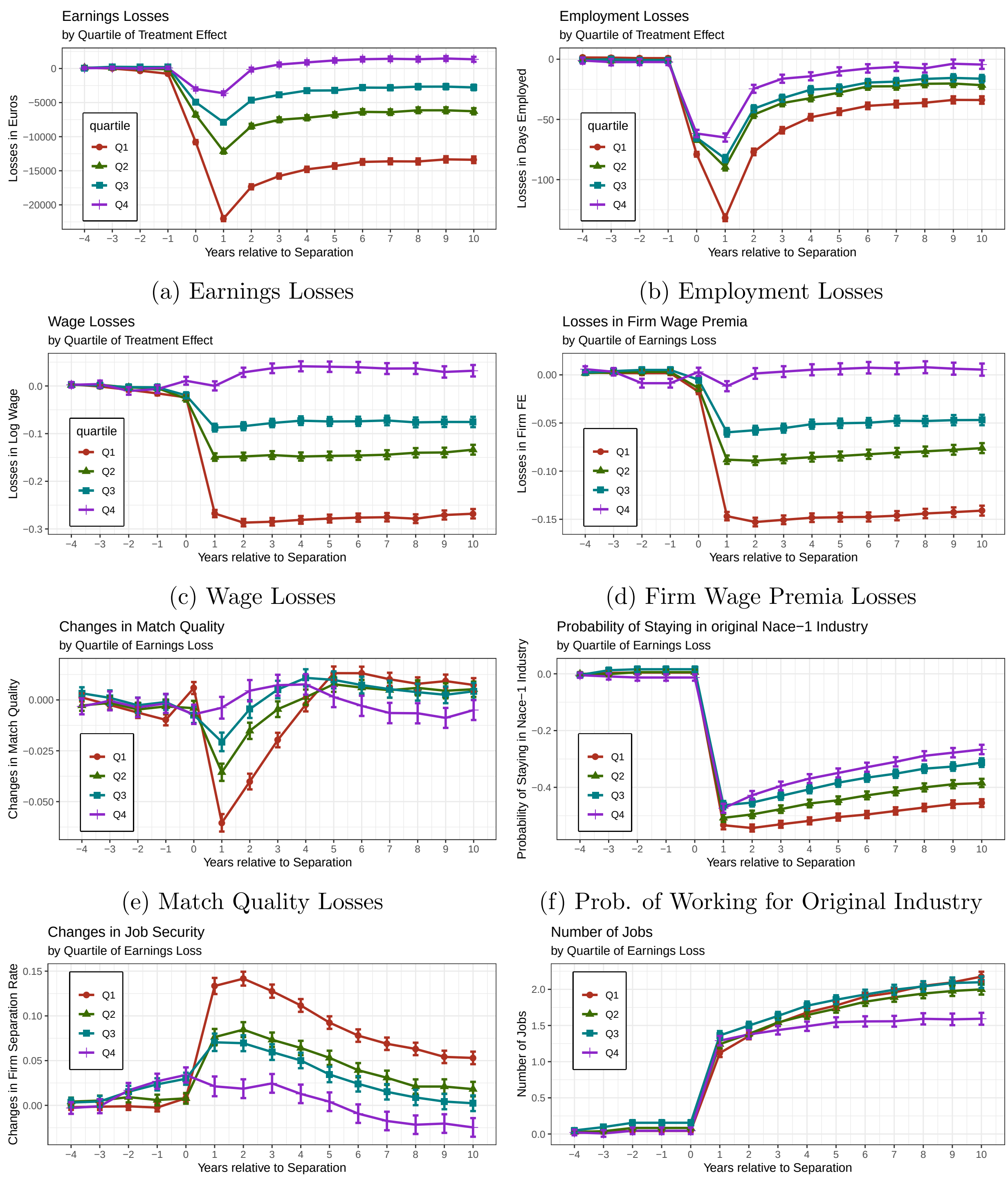
B. Post-Displacement Evolution

Why do workers face such different consequences from mass layoffs? Section VI will discuss in detail which *pre-displacement* worker and job characteristics are associated to higher losses. Understanding these is especially important, as these relationships can be used to effectively target labor market policies to workers vulnerable to large scarring effects. But before we turn to this, we study how labor market trajectories after the job loss differ for low versus high earnings loss individuals. We bin workers into quartiles according to their estimated earnings losses and estimate the event-study specification (2) for various left-hand side variables separately for every earnings loss quartile.

Figure 4 plots the dynamic evolution of earnings losses, employment losses and log-wage losses, and several job characteristics over time. The figure reconfirms that displaced workers experience vastly different labor market outcomes after job losses. The quarter of workers with the lowest losses (Q4) almost completely recover from employment losses and even gain in terms of earnings and wages on average in the long run. This is in stark contrast to the the quarter of workers with the highest losses (Q1), who face staggering wage declines of around 30%, and even ten years after the job loss are employed more than a month less, and earn 12,500€ less per year. This heterogeneity in the scarring effects is very persistence, with almost no convergence visible over the entire 10 year window after the job loss. The between-group difference in wage declines are essentially the same one year and ten years after the job loss.[19]

[18] Our groups are created based on the forest-based treatment effects as we find it the most natural. That said, other grouping criteria could be potentially used as well.

[19] More formal decomposition of earnings losses into wage and employment losses is presented in Section E



(a) Earnings Losses

(b) Employment Losses

(c) Wage Losses

(d) Firm Wage Premia Losses

(e) Match Quality Losses

(f) Prob. of Working for Original Industry

(g) Job Security Losses

(h) Number of Jobs

Figure 4: Earnings, employment and wage losses of displaced workers - event-study regression estimates, by quartile of estimated earnings losses as identified by a random forest. Period 0 corresponds to the separation year. Earnings and days employed are computed for the whole year, log-wages are computed as the log average daily wage from the employer on 1st January. Control group is selected via propensity score matching.

As already discussed above, Panel (d) shows that the changes in firm wage premia explain a large fraction of the changes in wages for every group. Both, wage and firm wage losses are measured in log points, thus by comparing the magnitudes we observe that the changes in the firm wage component explain between 50-60 percent of the overall change in wage losses. The rest is explained by changes in match quality, which are shown in Panel (e). Match quality is measured as the residual in the AKM Mincer wage regression Equation (4), and therefore captures all wage components which cannot be explained by an average year or age effect, nor by worker and firm fixed effects. The economic channels for losses in match quality are wide ranging, from losses of particularly good matches between the skill demands of the job and the skills offered by the worker, general human capital depreciation, or worsening of the bargaining position of workers after job loss to industry specific experience losses. Distinguishing these channels rigorously is difficult, but general human capital losses and worsening of the bargaining position are unlikely contributing factors to wage losses for the 50% of workers with the lowest losses. These workers either face match quality losses close to zero, or even gain, which is hard to square with human capital depreciation and lower bargaining positions after job losses. Some of the match quality losses might reflect that some workers are not able to return to their original industry and therefore lose the accumulated industry specific human capital. Panel (f) indeed confirms that high loss individuals have a lower propensity to stay in their original industry.

Panels (g) and (h) show in addition that workers with the highest losses move to firms with less job security and have more job changes consequently.

In summary, the effects of a mass layoff are far from uniform. Many workers recover in the long term or even gain in terms of wages, while a quarter of workers face permanent wage declines of around 30%. It is also notable how persistent the differences in the evolution of losses are. There is no visible convergence in any of the job characteristics. In most cases, the between group differences one year after the job losses are as large as ten years post-displacement. In the next section, we turn to the question of which pre-displacement characteristics are driving the heterogeneity.

VI. Sources of Earnings Losses

In this section we identify the *pre-displacement* characteristics which are the most strongly associated with higher earnings losses. This is insightful for multiple reasons. First, policy

in the Appendix. Nonetheless, changes in covariance between employment and wages are not very high, so the main conclusions are not very different from eyeball analysis of Panels (a)-(c) in Figure 4.

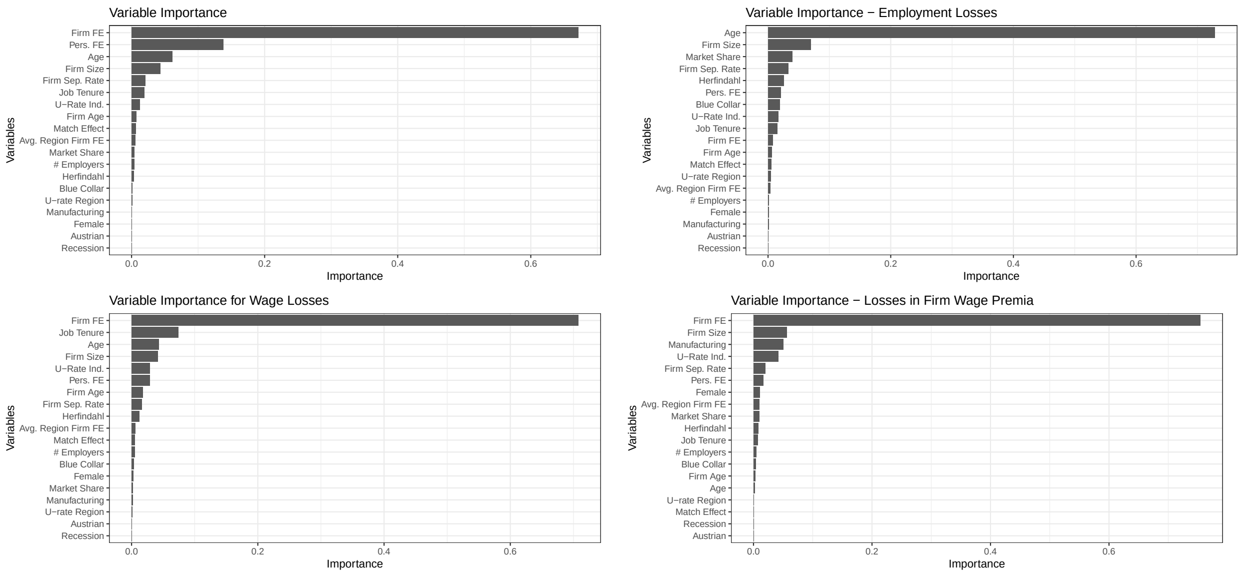


Figure 5: Variable frequency in splits in the GRF and the maximum depth level of nodes equal to 4. All values sum to 1.

makers can use this information to target specific programs aimed at alleviating the cost of job loss to high earnings loss individuals. Because we build separate forests for wage and employment losses, in principle even the type of program can be targeted. For example, job search assistance can be targeted to high employment loss individuals, whereas retraining or wage subsidy programs targeted to individuals predicted to face high wage losses. Second, identifying the most important pre-displacement characteristics determining the heterogeneity in earnings losses is informative about which factors are causing high earnings losses and thus helps to distinguish the many competing theories explaining earnings losses.

A. Horse Race Between Alternative Theories

Which variables are the most important to account for the heterogeneity of earnings losses? How do earnings losses vary by changing one factor at a time? Do these variables affect earnings losses through declines in wages, through losses in days employed, or both?

A compact and popular way in the machine learning literature to assess which factors are the most important is the occurrence frequency of variables in the splitting criteria. Variables chosen more frequently and earlier in the trees have a higher contribution in explaining the heterogeneity of scarring effects.[20] Figure 19 shows the variable importance measure for

[20] However, it is important to highlight that the raw version of this statistic without additional adjustments might be misleading because the greedy nature of tree-building prioritizes variables chosen first as more

losses in earnings, employment, wages, and firm wage premia. The striking result is that the pre-displacement firm fixed effect decile is by far the most important variable for explaining the heterogeneity in all consequences except for employment losses. The variable importance of firm fixed effects is close to 70 percent for earnings, wage and firm wage losses, while all other variables have variable important measures close or less than 10 percent. Only for employment losses the firm fixed effect does not play a major role. There, workers' age is the predominant factor with over 70 percent.

We also consider two alternative ways of judging the importance of different variables for explaining heterogeneous earnings losses. First, we carry out a standard variance decomposition of a linear projection of all variables onto individual costs identified through our forest. Section H in the Appendix describes the results in more detail, but the underlying conclusion is the same as before: Firm fixed effect is by far the most important variable explaining earnings, wage and firm wage losses, while workers' age is the most important factor for employment losses.

Another intuitive way to judge the importance of individual factors is to compute the elasticity of losses with respect to individual variables, holding all other factors constant, which we study in the next section.

B. Conditional Average Treatment Effects

We take advantage of our generalized random forests and we compute losses $\tau(\mathbf{z})$ conditional on various realizations of the partitioning variables, $\mathbf{z}$. This allows us to compute how losses in earnings, employment days, log-wage, and firm wage premia change with different values of one factor at a time, while holding all other variables fixed at their median. This way we are able to control for observable confounding factors. In addition, by comparing the outcomes for earnings, employment, and wage losses, we can study whether the channel affects earnings losses through employment or wage losses.

Figure 6 provides an overview by how much losses change, when moving individual variables from the first to the tenth decile, or from zero to one in case of dummies, while holding other variables constant at their median. The findings mirror the variable importance measures from before: earnings, wage and firm-wage losses change the most with firm fixed

crucial, undervaluing their true importance. For instance, consider two variables, z_1 (more important) and z_2. The algorithm selects z_1 first, making z_2 conditioned on z_1's outcomes. Consequently, z_1 appears in the split criteria less frequently than z_2, despite being more significant. To address this, depth-adjusted variable frequency is commonly used, weighting earlier splits more heavily. A decay exponent of -2, standard in machine learning, reduces the importance of split frequencies at each subsequent depth by 50%.

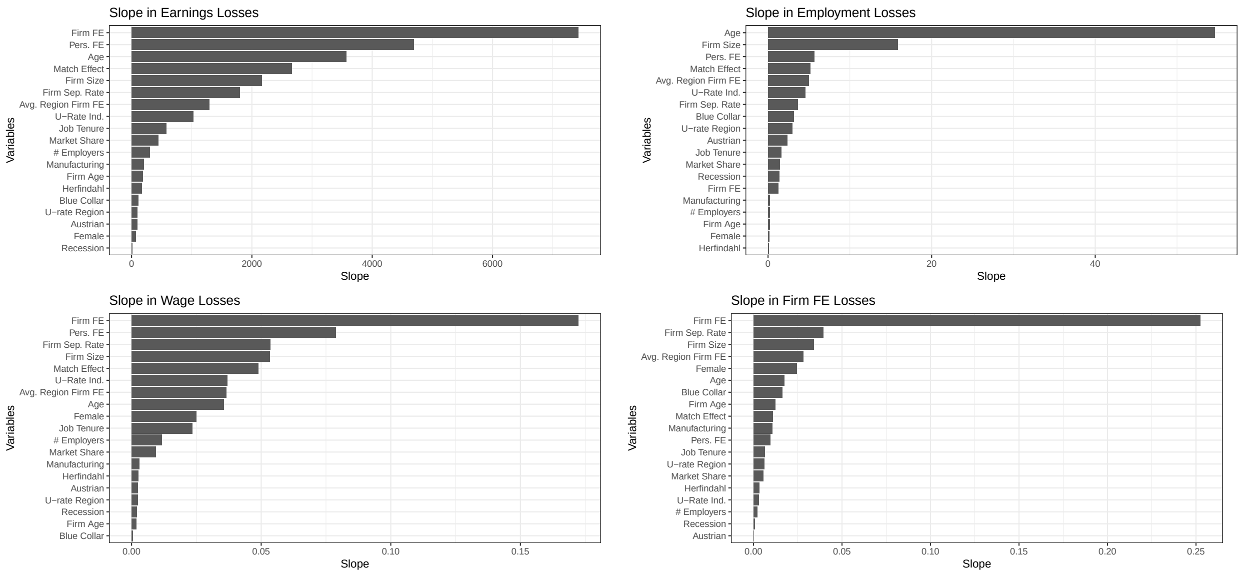


Figure 6: Partial effect of losses with respect to different variables. This illustrates the difference in losses between the highest and lowest decile (or 0/1 for binary variables) for each variable, while holding all other variables constant at their median values.

effects, and employment losses with worker's age. We discuss the most interesting cases in detail in the main text, while Figures 21 - 28 in the appendix report how earnings, employment, wage, and firm wage losses change with all the 19 different partitioning variables considered.

One potential criticism of the partial effects is that an individual with median characteristics might not be representative for the whole population. Those effects might be different for individuals with characteristics very different from the median. To tackle this critique, we use partial dependence plots proposed by Friedman (2001). The results are very close and can be found in in Appendix K.

B.1. Firm Wage Premia

Figure 7 shows the estimated consequences of job displacement, alongside 95% confidence intervals that account for the uncertainty arising from both, the machine learning procedure and the estimation procedure.[21] Since displaced workers and the selected control group might differ from the general population, we include a boxplot of the distribution of our sample over the variable of interest on top of every plot. The figure confirms the finding

[21]See Athey *et al.* (2019) and Sexton and Laake (2009) for a detailed description behind the estimation of standard errors.

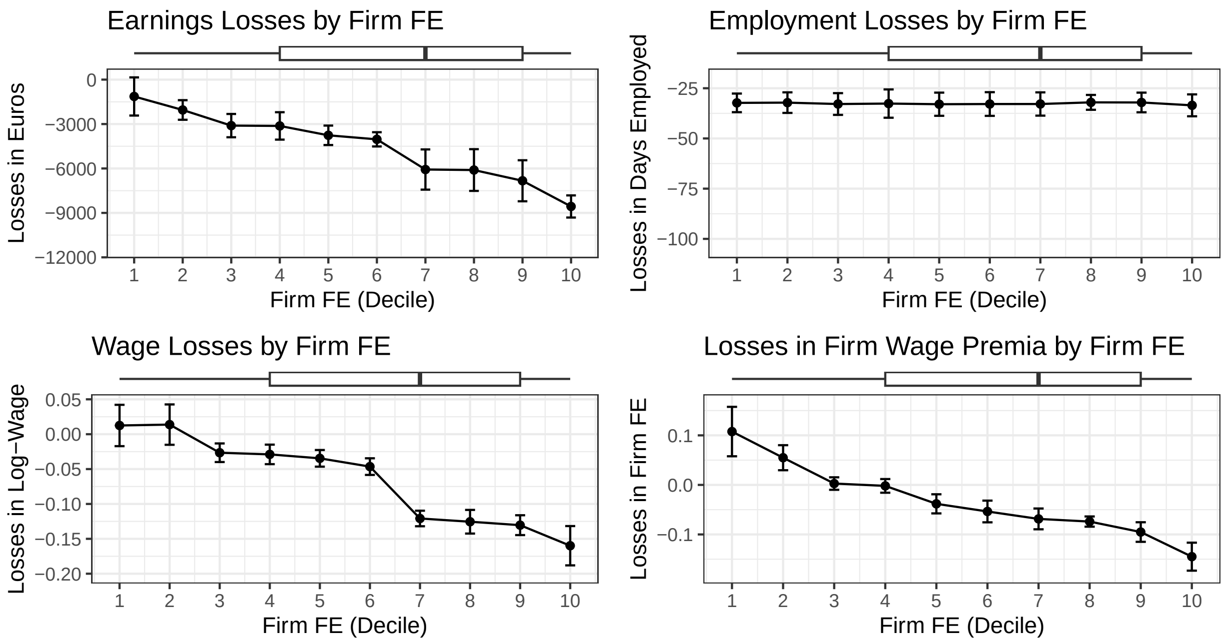


Figure 7: GRF estimates with 95% CI of losses in earnings, employment, wages, and firm premia by deciles of firm fixed effect. All other variables are set to their median values. The boxplots present the distribution of the partitioning variable in the dataset.

from the variable importance measure, that the displacement firm's wage premium is key to understand the cost of job loss. Out of all variables considered, earnings losses vary the most with firm wage premia. A worker separating from a firm paying in the bottom 10% of the firm pay distribution face earnings losses of €1,200, whereas workers formerly employed in the top decile paying jobs forgo more than 7 times as much, with earnings losses amounting to almost €8,500. These heterogeneous effects are also very precisely estimated. The confidence intervals for most declines are not wider than ±€600.

To understand whether this effect is purely coming from losses in employer specific wage components, or also through declines in employment, we study how firm fixed effects affect losses in employment, log-wages and firm fixed effects. It is visible, that the differences in earnings losses arise through wage losses, and not through employment losses. The differences in employment losses across firm FE deciles are small and statistically not significant. In contrast, the slope in wage losses mirrors the slope in earnings losses. Workers separating from the lowest paying firms do not face any wage losses, whereas wages decline by 16% for workers at the highest paying firms. The bottom right panel of Figure 7 reveals further striking results. First, losses in firm fixed effects, which are measured in log-wages, only explain part of wage losses. Wages across all firm fixed effect deciles decline by about 5 log-

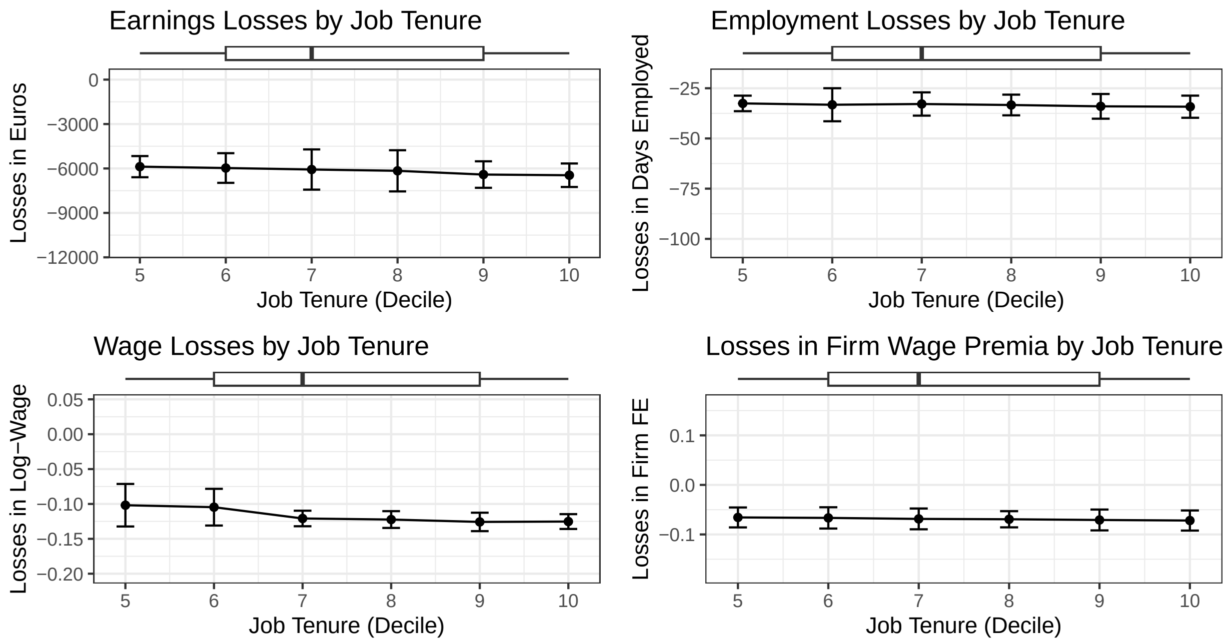


Figure 8: GRF estimates with 95% CI of losses in earnings, employment, wages, and firm premia by deciles of job tenure. All other variables are set to their median values. The boxplots present the distribution of the partitioning variable in the dataset.

points more than what can be explained by changes in firm fixed effects. But the *differences* in wage losses across firm fixed effect deciles are entirely explained by differences in lost firm wage premia. Second, changes in firm fixed effects show a mean reversion pattern. Workers employed in firms with above median firm pay face losses in firm wage premia, whereas workers employed in below median paying firms gain in terms of firm pay.

The importance of lost firm's wage premia in explaining earnings losses is further confirmed by studying how losses change by the availability of well-paying jobs in the region. As can be seen in Figure 26 in the Appendix, moving a worker with median characteristics from a region in the bottom decile of the average firm pay distribution to the highest decile reduces the estimated wage losses by 3.6 log points.

B.2. Job Specific and General Human Capital

Perhaps the most prominent theory about the sources of earnings losses is that workers lose either job specific or general human capital after job losses. We have included job tenure, which serves as a proxy for the firm specific human capital accumulated over the job tenure, and the pre-displacement worker fixed effects, which are a proxy for workers' transferable

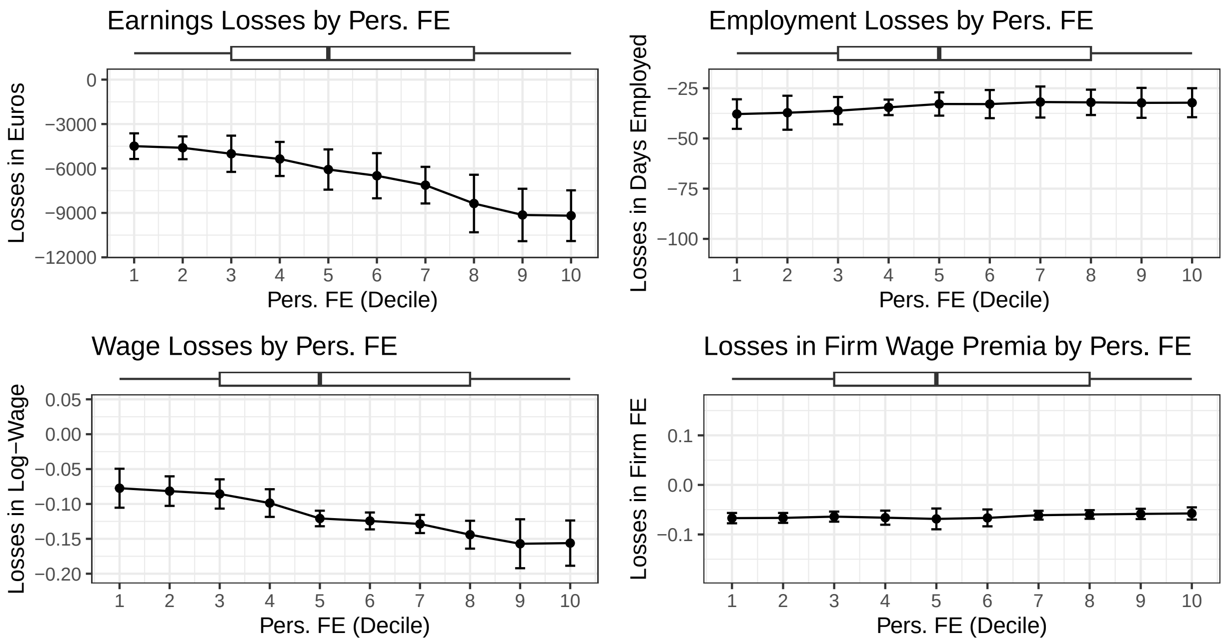


Figure 9: GRF estimates with 95% CI of losses in earnings, employment, wages, and firm premia by deciles of person fixed effect. All other variables are set to their median values. The boxplots present the distribution of the partitioning variable in the dataset.

skills. The hypothesis is that workers with more accumulated human capital are also more prone to face losses upon job displacement. Several papers have argued along this line and have shown that workers with higher tenure experience higher losses (*e.g.*, Topel, 1990; Jacobson *et al.*, 1993; Burdett *et al.*, 2020; Jarosch, 2023). This is in stark contrast to our findings. Figure 8 shows that neither earnings, employment, wage, nor firm wage losses are impacted much by changing job tenure, while holding other factors constant. This observation is very different from the one obtained using traditional sample splitting. Using a subgroup analysis we obtained a much higher slope in tenure, which would suggest that job-specific human capital is a very important force shaping overall losses. As we argue in subsequent Section VII, this discrepancy stems from the fact that traditional methods do not allow to vary one variable while keeping others fixed. This is of particular interest for job tenure as high-tenure workers are different in many other characteristics from low-tenure workers.

Through the lens of search and matching models, individuals with a high worker fixed effect, and thus high earning individuals might have accumulated higher savings and have more generous claims to unemployment benefits. This might allow them to be pickier at the job search, prolonging unemployment, but reducing wage losses, which is empirically docu-

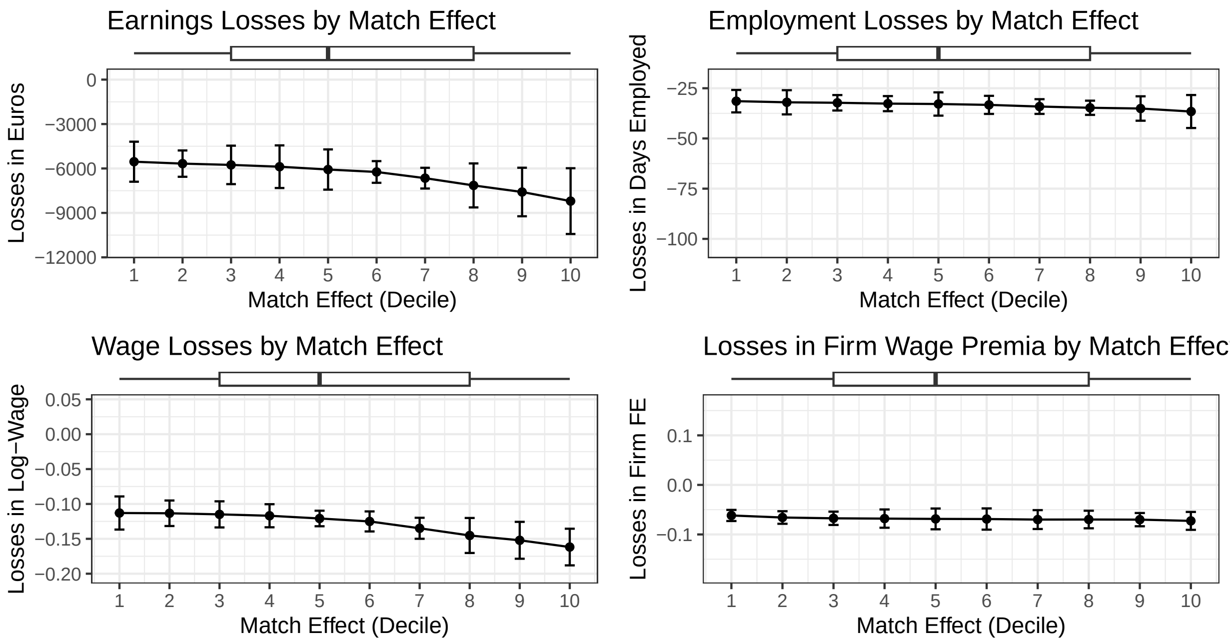


Figure 10: GRF estimates with 95% CI of losses in earnings, employment, wages, and firm premia by deciles of match quality. All other variables are set to their median values. The boxplots present the distribution of the partitioning variable in the dataset.

mented for Austria in Nekoei and Weber (2017). We find the opposite, Figure 9 shows that although a higher worker fixed effect marginally reduces employment losses, it is associated with much larger wage losses. A worker in the lowest decile faces wage losses of 7.5 percent, whereas the highest fixed effect workers experience losses around 15 percent, holding all other factors constant. Interestingly, these difference in wage losses are not driven by differences in firm wage losses, which suggests that indeed workers with previously high skills are more prone to skill losses.

B.3. Job Match Quality

Workers previously employed in particularly good matches could be prone to larger losses if they are not able to quickly find similarly good jobs. Figure 10 provides evidence in this direction. Workers with higher match specific wage components as captured by the residual term of equation (4) face overall higher losses. These higher losses are not driven by higher employment losses, because they do not change much with match effects. In contrast, workers displaced from a job in the lowest match quality deciles face wage losses of around 11 percent, whereas workers in the best matches face wage losses of 16 percent. This is

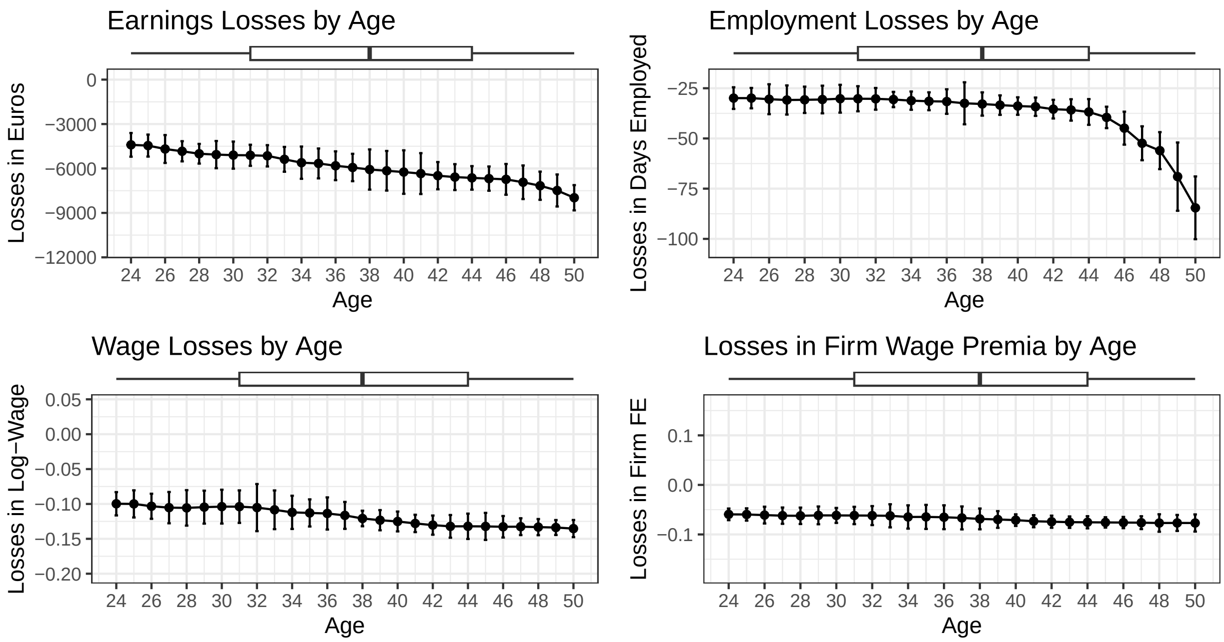


Figure 11: GRF estimates with 95% CI of losses in earnings, employment, wages, and firm premia by worker age. All other variables are set to their median values. The boxplots present the distribution of the partitioning variable in the dataset.

evidence that workers in particularly good matches fall off the match quality ladder as in Jovanovic (1979). However, a comparison of the slopes of losses for firm wage premia and match quality suggests that falling off the firm wage ladder is quantitatively the much bigger risk.

B.4. Workers' Age

Workers' age also plays an important role in understanding earnings losses. Figure 11 shows that total earnings losses increase nearly monotonically with age. By a large part, this is driven by higher employment losses for older workers, which face 3 times more lost employment days than workers below 35. Interestingly, wage losses are much less affected by age. All else equaly, wage losses increase from about 10 percent for young workers to 13 percent for old workers. These differences are not accounted for by firm wage premia, whose losses are flat over workers' age. Thus, it seems that older worker face more difficulties re-entering employment, but those who successfully find a job face similar wage losses to young workers.

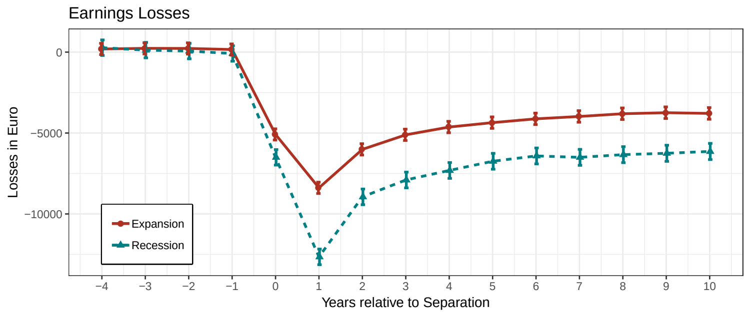


Figure 12: Earnings losses in recessions and expansions using the event-study specification (2). Observations were split into recessions and expansions categorized according to OECD definitions.

B.5. Cyclicality of Earnings Losses

How does the cost of job loss vary over the business cycle? A good understanding of the cyclicality of losses is a premise for designing good labor-market policies. Previously, this question was studied for example by Davis and Von Wachter (2011). Using a traditional subgroup analysis, they documented substantial differences in losses across workers laid off during downturns and expansions. However, as already argued, sample splitting has the drawback that it does not allow to keep confounding factors fixed. Consequently, in this particular context one cannot disentangle between the direct effect associated with the recession itself and the compositional effect linked to time-varying characteristics of displaced workers. As we discuss below, our machine-learning methodology is well tailored to identify those two components separately.

We study how losses change with whether the Austrian economy is in a recession. As can be seen in Figure 6, earnings losses do not vary much with the recession indicator, when holding confounding factors constant. For example, a displaced worker faces very similar earnings losses than an observationally identical worker displaced during an expansion. Also the variable importance associated with the state of the local labor market and industry of workers is very limited.[22] In contrast, if we split our sample and re-estimate earnings losses using the event-study specification (2) separately for workers separating in recessions and expansions we find significantly higher losses in recessions, as can be seen in Figure 12. This finding is consistent with similar observations documented by Davis and Von Wachter

[22]Earnings losses almost do not change in the regional unemployment rate. The impact of the industrial unemployment rate is for most values nearly the same except for bottom deciles where the treatment effects feature higher standard errors.

(2011) and Schmieder *et al.* (2023) for the United States and Germany. How can those seemingly contradictory observations be reconciled? One must bear in mind that there are two sources accounting for the variation of earnings losses across the business cycle observed in Figure 12: the impact of a recession per se and the compositional differences due to the fact that different workers are displaced during a recession than during an expansion. In particular, the latter component can be relevant as Mueller (2017) documented that the composition of unemployed workers changes significantly over the business cycle. The cyclicality of earnings losses can thus be decomposed into two components: the pure recession effect and compositional differences. Formally, this decomposition can be expressed as follows:

$$\begin{aligned}\int \tau(\mathbf{z}|rec=1)\mathrm{d}F(\mathbf{z}|rec=1) - \int \tau(\mathbf{z}|rec=0)\mathrm{d}F(\mathbf{z}|rec=0) \quad &= \\ \underbrace{\int \left[\tau(\mathbf{z}|rec=1) - \tau(\mathbf{z}|rec=0)\right] \mathrm{d}F(\mathbf{z}|rec=1)}_{\text{Recession effect}} &+ \\ \underbrace{\int \tau(\mathbf{z}|rec=0)\mathrm{d}F(\mathbf{z}|rec=1) - \int \tau(\mathbf{z}|rec=0)\mathrm{d}F(\mathbf{z}|rec=0)}_{\text{Compositional difference}}&, \end{aligned} \tag{8}$$

where $F(\mathbf{z}|rec=r)$ denotes the distribution of worker and job characteristics, $\mathbf{z}$, conditioned on the aggregate state of the economy, with $rec=1$ during a recession and 0 otherwise. Then $\int \tau(\mathbf{z}|rec=r)\mathrm{d}F(\mathbf{z}|rec=r)$ represents the average treatment effect of job loss for all displaced workers conditional on the aggregate state $rec=r$. In our application, these values can be approximated with $\frac{1}{N_r}\sum_{i=1}^{N_r} \mathbb{1}_{\{rec_i=r\}}\hat{\tau}(\mathbf{z}_i)$, where N_r is the number of displaced workers when $rec=r$.

Equipped with our random forest, we can identify the counterfactual treatment effect of job loss in expansions for workers displaced in recessions. The average difference between the cost of job loss in a recession versus in an expansion for identical workers represents the pure effect of a recession, as embodied by the first RHS element of the decomposition in Equation (8). Compositional differences are captured by the latter element. While a traditional sample-splitting approach captures only the overall effect (i.e., the LHS of Equation (8)), our random forest allows us to identify the two components, the recession effect

	(1)	(2)	(3)	(4)	(5)
		Recession Effect		Composition	
	Difference	Level	Share	Level	Share
Recession dist.: Eq. (8)	-569.54	-42.18	0.07	-527.36	0.93
Expansion dist.: Eq. (9)	-569.54	-34.22	0.06	-535.32	0.94

Table 3: Decomposition of the difference in losses during recessions and expansions into a recession effect and a composition effect. See text for details.

and compositional differences, separately.[23] Table 3 shows that the recession impact by itself is very small and accounts for only 6-7% of the overall differences in earnings losses and the remaining part stems from compositional changes of displaced workers. If we break down individuals into those displaced during expansion and recession (Table 11), it turns out that workers displaced in bad times exhibit higher job tenure and, prior to the job termination, they were hired at smaller firms with slightly higher firm wage premia. Thus, our finding is that, if anything, policies should not be recession-dependent, but rather condition on job characteristics of displaced workers.

B.6. Other Factors

Of the remaining factors in Figure 6, firm size and the firm separation rate have the strongest impact. Firm size is in fact the second most important variable for employment losses and the fourth most important for wage losses. Workers separating from the largest decile of firms face employment losses of additional 15 days per year compared to the smallest firms, and conditional on reemployment face additional wage declines of more than 5 percent.

[23] Alternatively, we can use the distribution of the workers displaced in expansion as the baseline:

$$\begin{aligned}\int \tau(\mathbf{z}|rec=1)\mathrm{d}F(\mathbf{z}|rec=1) - \int \tau(\mathbf{z}|rec=0)\mathrm{d}F(\mathbf{z}|rec=0) \quad &= \\ \underbrace{\int \left[\tau(\mathbf{z}|rec=1) - \tau(\mathbf{z}|rec=0)\right] \mathrm{d}F(\mathbf{z}|rec=0)}_{\text{Recession effect}} &+ \\ \underbrace{\int \tau(\mathbf{z}|rec=1)\mathrm{d}F(\mathbf{z}|rec=1) - \int \tau(\mathbf{z}|rec=1)\mathrm{d}F(\mathbf{z}|rec=0)}_{\text{Compositional difference}}&. \end{aligned} \tag{9}$$

We report results for both decompositions in Table 3.

The higher losses in large firms is consistent with spillovers through local demand effects (Gathmann *et al.*, 2020). Workers displaced from firms with historically high turnover rates face overall lower losses. This naturally operates through the control group, which is more prone to job losses themselves in high turnover firms. Interestingly, the historic firm level separation rate do not affect employment losses much, but is the second most important factor for firm wage losses, and the third most important for wage losses, suggesting that the control group in high turnover firms are also more prone to fall down the firm wage ladder. Thus, the displacement costs are especially elevated due to losses in well-paying jobs with high job security, consistent with the findings of Jarosch (2023).

Changing any of the other factors do not move employment losses by more than three to four days per year and wage losses by more than three percent. Although these other factors do not play an important role, they are interesting nevertheless as many are in contrast to conventional wisdom and empirical evidence based mostly on sample splitting. For example, Helm *et al.* (2023) argue that manufacturing workers are especially hurt by job displacements, as these jobs are on a secular decline. In contrast, we find that conditional on confounding factors, manufacturing workers do not face elevated losses. Although workers in manufacturing also face substantially higher losses in Austria if we split our sample, once we hold confounding factors constant, it does not play a role anymore whether a service or a manufacturing worker is displaced. Thus, in the context of Austria, the higher losses in manufacturing purely arise because these jobs are typically in higher paying firms.

Gender differences in the cost of job loss recently gained attention in the literature, Illing *et al.* (2024) find large gender difference in losses for Germany. In cotrast we find rather small differences. Ceteris paribus, there is essentially no difference in employment losses, and a 2.5 log point higher wage losses for females, which is entirely explained by moving to lower paying firms.

Surprisingly, the unemployment rate in the region, as well as the unemployment rate in the displacing industry barely have an effect on the magnitude of losses. As with the recession indicator, it seems that what matters more for the understanding of earnings losses is especially how well the firm was paying, not so much economic circumstances during the job loss.

We do not find that losses change much with indicators for being blue collar or Austrian, with past labor market mobility, and with the region/industry Herfindahl labor market concentration index, firm age or labor market share.

VII. Heterogeneity Detection: Comparison with Other Methods

In this section, we compare earnings losses identified through our generalized random forest with other methods commonly employed in empirical research. In this exercise we consider two views on treatment heterogeneity: (*i*) the observable one, where the heterogeneity is modeled as a function of some variables and (*ii*) latent one where treatment effects are independent of any characteristics. For observable heterogeneity, we adopt two approaches: (*i*) splitting our sample based on several individual characteristics and estimating the DiD regression (1) for each group separately; (*ii*) estimating the DiD regression where the post-displacement dummy for the treatment group interacts with other observables. On the other hand, the latent heterogeneity is studied by employing (i) Changes-in-Changes and (ii) quantile DiD, both proposed by Athey and Imbens (2006).[24]

A. Conditional Treatment Effects

We repeat the DiD analysis, this time splitting our sample based on the values of each variable used to grow our random forest. Figure 13 presents the estimated average annual earnings losses for selected variables.[25] The impact of firm wage premia appears quite consistent across both approaches. However, sample splitting also spuriously detects other variables as significant drivers of earnings losses. Notably, job tenure — a widely discussed factor in the literature as the most prominent source of earnings losses (*cf.*, Topel, 1990; Jacobson *et al.*, 1993; Burdett *et al.*, 2020; Jarosch, 2023) — is portrayed as having more influence than it may actually have in reality. Unlike our algorithmic approach, which upholds the *ceteris paribus* condition by varying only one variable at a time, sample splitting fails to isolate single factor effects, potentially leading to detecting spurious relationships.[26] For instance, firm wage premia are generally higher in the manufacturing sector. Hence, individuals segmented by sector inadvertently reflect different wage premia levels, muddling the drivers of heterogeneity in earnings losses shown in Figure 13. Is it due to firm wage premia or sectoral differences?

[24] All those methods are popular in the empirical literature. For instance, D'Acunto *et al.* (2022) use subgroup analysis for studying heterogeneous impact of unconventional fiscal policy and forward guidance on private consumption. Havnes and Mogstad (2015) use, among others, quantile DiD and Changes-in-Changes to evaluate the distributive impact of a child-care reform on future earnings in Norway. Moreover, quantile regression, in the context of analyzing earnings losses, was recently employed by Jarosch (2023, Section 4.1.4).

[25] Results for all variables used in the partitioning are provided in Figures 45 to 47 of Appendix M.

[26] Figure 16 indicates correlations among variables, suggesting possible distortions.

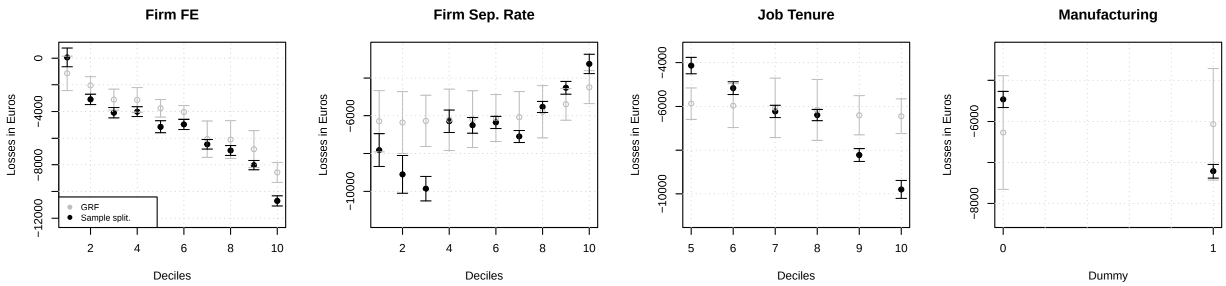


Figure 13: Earnings losses by different values of four selected variables: firm wage premia, job tenure, firm separation rate, and manufacturing indicator. Black and grey dots show earnings losses implied by sample splitting and the random forest, respectively. The results for all partitioning variables are relegated to Figures 45–47.

Subgroup analysis does not provide a definitive answer. On the other hand, our algorithmic method isolates the impact of firm wage premia and sector, demonstrating that the former significantly affects earnings losses, while the latter does not at all.[27]

To tackle the issue of covarying observables, one strategy involves employing a regression model that interacts the treatment for the treated group with certain individual characteristics. In this exercise, we estimate a model that balances flexibility with computational feasibility. We fully interact each variable's values with those of all other variables, considering up to second-order interaction effects. This approach remains agnostic about the importance of specific variables.[28] Formally, our regression takes the following form:

$$y_{it} = \mathbb{1}(t \geq t^*) \times D_i \left(\sum_j \sum_{p \in v(j)} \sum_{k \neq j} \sum_{q \in v(k)} \tau(X_{ij}^p, X_{ik}^q) X_{ij}^p \times X_{ik}^q \right) + \theta D_i + \gamma_t + \epsilon_{it}, \qquad (10)$$

where X_{im}^r denotes the dummy variable for worker i's characteristic m associated with value r, and $v(m)$ denotes the set of possible values for characteristic m. Then, $\tau(X_{ij}^p, X_{ik}^q)$ represents the (partial) treatment effect for observations where variable j takes value p *and* variable k takes value q.[29] While this way we obtain a quite flexible functional specifica-

[27] A minor comment regarding standard errors for both methods is worth mentioning. Admittedly, the errors are lower for subgroup analysis than for the random forest. Nonetheless, by no means they are comparable. The reason for this is that the random-forest standard errors in addition to estimation uncertainty incorporates uncertainty caused by the model specification. In the subgroup analysis the model specification is assumed to be known since the very beginning.

[28] While the flexibility of this approach is still lower than in our random forest, it surpasses the flexibility of models where treatment is interacted with a single variable in a predetermined functional form.

[29] The overall treatment effect is the sum of all estimated coefficients corresponding to the dummies relevant

tion that allows us to consider many possible patterns of heterogeneity in earnings losses, a very high number of identified parameters (15,944) might result in much lower estimation accuracy. Indeed, this is the case.

	Treatment Percentiles			Fraction of individuals with losses *statistically significant*:		
	25%	50%	75%	below the median	above the median	below or above the median
Quadratic Model	-14,309	-6,177	1,400	0.18	0.17	0.35
GRF	-8,339	-5,768	-3,491	0.32	0.35	0.67

Table 4: Heterogeneity in losses and its estimation precision. The left panel of the table shows the bottom quartile, the median, and the top quartile of earnings losses for both, the quadratic model and the random forest estimates. The right panel shows the fraction of workers facing losses below the median level in a statistically significant way ($\tau(x_i) + 1.96 \cdot se(x_i) < \mu$), or losses above the median level in a statistically significant way ($\tau(x_i) - 1.96 \cdot se(x_i) > \mu$), or either of both cases is true.

As Table 4 shows, our random forest and the quadratic model from Equation (10) identify the median earnings losses at the level of around €6,000 per year. The very high flexibility of the quadratic model is confirmed by the fact that the interquartile range of treatment effects from the quadratic model is larger than for the random forest. That being said, only for 35% individuals the earnings losses are significantly different from the median estimate, which suggests that a large portion of the identified heterogeneity may be caused by estimation noise. On the other hand, in the random forest 67% individuals exhibit earnings losses significantly different from the median level.[30]

B. Latent heterogeneity

Variation in earnings losses, independent of observed characteristics, offers an alternative approach to conditional average treatment effects. In this exercise, we use two non-linear extensions of the DiD method: Changes-in-Changes (hereinafter CiC) and Quantile Difference-in-Difference (hereinafter QDiD). Both methods enable the identification of changes in the distribution of the dependent variable caused by treatment on the treatment group.[31] The

for an individual.

[30]It is noteworthy to reiterate that, similarly to sample splitting, the standard errors from the quadratic model do not include specification uncertainty like in the random forest. Then, the much higher precision of the random forest is even more striking.

[31]A detailed discussion on how to impute the comparison group in the DiD setting can be found in Athey and Imbens (2006), where CiC and QDiD were first proposed.

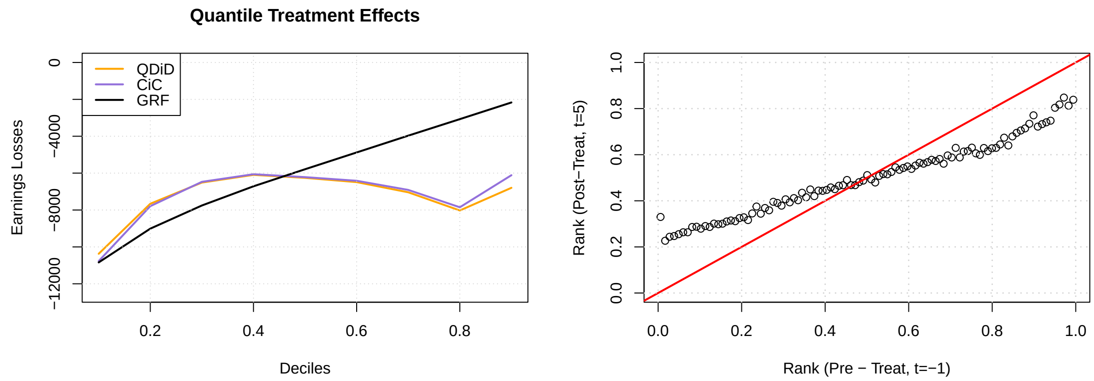


Figure 14: The left panel shows changes in deciles of earnings after and before treatment computed with QDiD and CiC. For GRF, we computed deciles of individual treatment effects. The right panel depicts the relationship between earnings quantile before displacement and the average earnings quantile 5 years after job separation. The quantile statistics are computed for the treatment group. The slope of the dashed red line is equal to 45°.

treatment effect for a specific q-quantile is identified simply as the difference between the q-quantile of earnings in the treatment group and the q-quantile in the control group. These procedures enable the evaluation of overall distributional responses to the treatment but require strong assumptions for identifying individual-level treatment effects. Specifically, quantile regressions can only identify individual earnings losses if the ranks of earnings before and after the treatment are perfectly correlated. Yet, the right panel of Figure 14 illustrates that this condition is not met.[32] Workers who were earning above the median before displacement on average fall in the earnings ranking 5 years after job separation, while individuals' ranks from the bottom 50% increase. Spearman's rank correlation of earnings quantiles between both periods is equal to only 0.54 and the average distance between pre- and post-displacement rank is equal to 20 quantile points. The left panel of Figure 14 shows that the distribution of individual earnings losses is much more dispersed than for distributions identified through non-linear DiD. This is caused by the fact mentioned above – different workers fell off the earning ladder in a different way, which changes their earnings rank after job separation. This results in higher dispersion in earnings losses in comparison to losses implied by QDiD and CiC. The presented exercise shows that our random-forest procedure requires much less restrictive assumptions to identify individual earnings losses

[32]In particular, this leads to non-monotone estimates for quantile effects, as shown in the left panel of Figure 14. It is caused by the fact that the difference between the 80th quantile of the treatment group and the control group is lower than the difference observed at lower quantiles.

and that the variation in individual losses implied from quantile regressions are severely underestimated.

VIII. Conclusions

We implement a generalized random forest (Athey *et al.*, 2019) to a DiD setting to study the sources of earnings losses of displaced workers. This methodology allows us to make a number of important empirical contributions to the existing literature. First, we document the heterogeneity in the causal cost of job loss across individuals. Using the universe of Austrian social security records from 1984 through 2019, we show that there is substantial heterogeneity in earnings losses across individuals. We document that a quarter of workers face wage losses of 30%, whereas 25% of workers face no losses or even gain. Second, the machine learning procedure allows us to conduct a horse race between many competing theories about earnings losses, while controlling for observable confounding factors. We find that the pre-displacement firm wage premium is by far the most important channel for earnings, wage, and firm wage losses, while workers' age is the most important for the level of employment losses. Holding all other variables fixed, workers separating from the lowest quintile of the firm pay distribution experience modest wage gains from job displacement, whereas workers separating from the best paying decile of firms face wage losses of 16 log points.

In addition, we use the fact that our methodology enables us to estimate earnings losses at the individual level and decompose the cyclical variation of earnings losses into a pure recession effect and compositional differences due to the fact that different workers are displaced during a recession than during an expansion. During recessions, the composition of displaced workers shifts towards worker and job characteristics that are associated with higher losses, which explains over 90% of the cyclicality. This highlights the importance of the ability of our machine learning approach to hold worker and job characteristics constant when studying the impact of different channels. All in all, our findings provide evidence that earnings losses can be understood by mean reversion in firm wage premia, rather than by a destruction of firm-specific human capital, while earnings losses for older workers are mostly driven by employment losses.

We compare the heterogeneity of earnings losses identified through our method with traditional empirical approaches: interacting the treatment variable with others and quantile regression. Although DiD regression with multiple interaction terms detects heterogeneity, it is predominantly influenced by estimation noise. Furthermore, quantile regression estimates offer limited insights into individual earnings losses due to the imperfect correlation of earnings ranks before and after treatment.

Our paper documents that the consequences of job loss are far from uniform. A clear policy implication of our findings is that policies aimed to mitigate the consequences of job loss such as firm bailouts, employment protection, and employment subsidies such as short time work schemes, which are often applied in a non-discriminatory fashion, should be, if anything, targeted. Policy makers could use our results to target specific programs aimed at alleviating the cost of job loss to high earnings loss individuals. Because we build separate forests for wage and employment losses, in principle even the type of program can be targeted. For example, job search assistance can be targeted to high employment loss individuals, whereas retraining or wage subsidy programs targeted to individuals predicted to face high wage losses.

References

Abowd, J. M., Kramarz, F. and Margolis, D. N. (1999). High wage workers and high wage firms. *Econometrica*, **67** (2), 251–333.

Athey, S. and Imbens, G. W. (2006). Identification and inference in nonlinear difference-in-differences models. *Econometrica*, **74** (2), 431–497.

—, Simon, L. K., Skans, O. N., Vikstrom, J. and Yakymovych, Y. (2023). The heterogeneous earnings impact of job loss across workers, establishments, and markets. *arXiv preprint arXiv:2307.06684*.

—, Tibshirani, J. and Wager, S. (2019). Generalized random forests. *The Annals of Statistics*, **47** (2), 1148–1178.

Bertheau, A., Acabbi, E. M., Barceló, C., Gulyas, A., Lombardi, S. and Saggio, R. (2023). The unequal consequences of job loss across countries. *American Economic Review: Insights*, **5** (3), 393–408.

Breiman, L. (2001). Random forests. *Machine Learning*, **45** (1), 5–32.

—, Friedman, J. H., Olshen, R. A. and Stone, C. J. (1984). *Classification and Regression Trees*. Statistics/Probability Series, Belmont, California, U.S.A.: Wadsworth Publishing Company.

Burdett, K., Carrillo-Tudela, C. and Coles, M. (2020). The Cost of Job Loss. *The Review of Economic Studies*, **87** (4), 1757–1798.

Cahuc, P., Postel-Vinay, F. and Robin, J.-M. (2006). Wage bargaining with on-the-job search: Theory and evidence. *Econometrica*, **74** (2), 323–364.

Card, D., Heining, J. and Kline, P. (2013). Workplace heterogeneity and the rise of west german wage inequality*. *The Quarterly Journal of Economics*, **128** (3), 967–1015.

Couch, K. A. and Placzek, D. W. (2010). Earnings losses of displaced workers revisited. *The American Economic Review*, pp. 572–589.

Davis, J. M. and Heller, S. B. (2020). Rethinking the benefits of youth employment programs: The heterogeneous effects of summer jobs. *Review of economics and statistics*, **102** (4), 664–677.

Davis, S. J. and Von Wachter, T. (2011). Recessions and the costs of job loss. *Brookings Papers on Economic Activity*, **43** (2 (Fall)), 1–72.

D'Acunto, F., Hoang, D. and Weber, M. (2022). Managing Households' Expectations with Unconventional Policies. *The Review of Financial Studies*, **35** (4), 1597–1642.

Farber, H. S. (2011). *Job loss in the Great Recession: Historical perspective from the displaced workers survey, 1984-2010*. Tech. rep., National Bureau of Economic Research.

— (2017). Employment, hours, and earnings consequences of job loss: Us evidence from the displaced workers survey. *Journal of Labor Economics*, **35** (S1), S235–S272.

Friedman, J. H. (2001). Greedy function approximation: A gradient boosting machine. *The Annals of Statistics*, **29** (5), 1189–1232.

Gathmann, C., Helm, I. and Schönberg, U. (2020). Spillover effects of mass layoffs. *Journal of the European Economic Association*, **18** (1), 427–468.

Gregory, V., Menzio, G. and Wiczer, D. G. (2021). *The alpha beta gamma of the labor market*. Tech. rep., National Bureau of Economic Research.

Halla, M., Schmieder, J. and Weber, A. (2018). Job displacement, family dynamic, and spousal labor supply.

Havnes, T. and Mogstad, M. (2015). Is universal child care leveling the playing field? *Journal of Public Economics*, **127** (C), 100–114.

Helm, I., Kügler, A. and Schönberg, U. (2023). Displacement effects in manufacturing and structural change.

Hethey-Maier, T. and Schmieder, J. F. (2013). Does the use of worker flows improve the analysis of establishment turnover? evidence from german administrative data. *Schmollers Jahrbuch: Journal of Applied Social Science Studies/Zeitschrift für Wirtschafts-und Sozialwissenschaften*, **133** (4), 477–510.

Illing, H., Schmieder, J. and Trenkle, S. (2024). The gender gap in earnings losses after job displacement. *Journal of the European Economic Association*, p. jvae019.

Jacobson, L., LaLonde, R. and Sullivan, D. (1993). Earnings losses of displaced workers. *The American Economic Review*, pp. 685–709.

Jarosch, G. (2023). Searching for job security and the consequences of job loss. *Econometrica*, **91** (3), 903–942.

Jovanovic, B. (1979). Job matching and the theory of turnover. *Journal of Political Economy*, **87** (5), 972–990.

Knaus, M. C., Lechner, M. and Strittmatter, A. (2022). Heterogeneous employment effects of job search programs: A machine learning approach. *Journal of Human Resources*, **57** (2), 597–636.

Lachowska, M., Mas, A. and Woodbury, S. A. (2020). Sources of displaced workers' long-term earnings losses. *American Economic Review*, **110** (10), 3231–3266.

Lindo, J. M. (2011). Parental job loss and infant health. *Journal of Health Economics*, **30** (5), 869–879.

McCall, J. (1970). Economics of information and job search. *The Quarterly Journal of Economics*, pp. 113–126.

Mueller, A. I. (2017). Separations, sorting, and cyclical unemployment. *American Economic Review*, **107** (7), 2081–2107.

Neal, D. (1995). Industry-specific human capital: Evidence from displaced workers. *Journal of Labor Economics*, **13** (4), pp. 653–677.

Nekoei, A. and Weber, A. (2017). Does extending unemployment benefits improve job quality? *American Economic Review*, **107** (2), 527–61.

Rege, M., Telle, K. and Votruba, M. (2011). Parental job loss and children's school performance. *The Review of Economic Studies*, **78** (4), 1462–1489.

Schaller, J. and Stevens, A. H. (2015). Short-run effects of job loss on health conditions, health insurance, and health care utilization. *Journal of health economics*, **43**, 190–203.

Schmieder, J. F., Von Wachter, T. and Heining, J. (2023). The costs of job displacement over the business cycle and its sources: evidence from germany. *American Economic Review*, **113** (5), 1208–1254.

Sexton, J. and Laake, P. (2009). Standard errors for bagged and random forest estimators. *Computational Statistics and Data Analysis*, **53** (3), 801 – 811, computational Statistics within Clinical Research.

Sullivan, D. and Von Wachter, T. (2009). Job displacement and mortality: An analysis using administrative data. *The Quarterly Journal of Economics*, **124** (3), 1265–1306.

Topel, R. (1990). Specific capital and unemployment: Measuring the costs and consequences of job loss. In *Carnegie-Rochester Conference Series on Public Policy*, Elsevier, vol. 33, pp. 181–214.

Zweimüller, J., Winter-Ebmer, R., Lalive, R., Kuhn, A., Wuellrich, J.-P., Ruf, O. and Büchi, S. (2009). Austrian social security database. *Available at SSRN 1399350*.

Online Appendix

A. Propensity Score Matching

The table below shows the outcomes of the propensity score matching for each year separetly. Group 1 refers to displaced workers, group 2 to the control group and group 3 to unselected workers. The left panels of Figure 15 show the evolution of earnings, days employed and log-wages in the treatment and control group.

group	Age	$\log(w_{t-1})$	$\log(w_{t-2})$	Job Tenure	Manufact.	Firm Size	Female	propensity	N
Year = 1989									
1	37.74	4.46	4.45	1604.08	0.55	199.06	0.35	50.11	1597
2	37.68	4.45	4.44	1593.78	0.60	207.23	0.36	50.11	1597
3	37.74	4.49	4.46	1638.85	0.47	886.00	0.37	50.07	543851
Year = 1990									
1	37.12	4.33	4.31	1797.71	0.55	126.87	0.50	50.20	1971
2	36.97	4.35	4.33	1805.58	0.57	123.89	0.47	50.20	1971
3	37.64	4.50	4.47	1862.16	0.47	856.58	0.37	50.09	533558
Year = 1991									
1	36.96	4.35	4.32	1992.85	0.48	141.36	0.47	50.19	2143
2	37.07	4.34	4.31	1987.76	0.49	145.07	0.50	50.19	2143
3	37.53	4.53	4.49	2016.12	0.48	985.58	0.37	50.10	543431
Year = 1992									
1	36.95	4.37	4.33	2034.17	0.55	194.93	0.49	50.28	3439
2	37.24	4.36	4.32	2029.65	0.56	205.67	0.51	50.28	3439
3	37.35	4.55	4.51	2143.80	0.48	968.57	0.36	50.15	553273
Year = 1993									
1	36.79	4.47	4.46	2215.43	0.72	169.46	0.38	50.40	4550
2	36.79	4.45	4.44	2214.94	0.72	169.52	0.39	50.40	4550
3	37.23	4.55	4.53	2253.20	0.48	958.74	0.36	50.19	565187
Year = 1994									
1	37.14	4.40	4.39	2203.21	0.56	85.15	0.44	50.28	2354
2	37.09	4.41	4.39	2176.10	0.54	83.05	0.44	50.28	2354
3	37.11	4.56	4.53	2341.56	0.47	932.61	0.36	50.10	574710
Year = 1995									
1	37.87	4.43	4.40	2368.50	0.44	260.58	0.49	50.29	4873

(continued)

group	Age	$\log(w_{t-1})$	$\log(w_{t-2})$	Job Tenure	Manufact.	Firm Size	Female	propensity	N
2	37.69	4.44	4.40	2372.35	0.44	254.81	0.49	50.29	4873
3	37.15	4.57	4.54	2441.55	0.47	899.02	0.36	50.21	576684
Year = 1996									
1	37.65	4.52	4.50	2296.31	0.65	201.44	0.36	50.35	4733
2	37.43	4.51	4.49	2257.01	0.68	213.49	0.36	50.35	4733
3	37.37	4.59	4.55	2553.28	0.45	889.40	0.37	50.20	579847
Year = 1997									
1	37.74	4.52	4.51	2456.51	0.54	90.10	0.34	50.17	1707
2	37.71	4.52	4.51	2529.01	0.52	88.62	0.35	50.17	1707
3	37.42	4.58	4.57	2623.88	0.45	859.78	0.37	50.07	592559
Year = 1998									
1	38.03	4.50	4.49	2455.48	0.44	188.51	0.42	50.15	2395
2	38.33	4.50	4.48	2405.43	0.44	193.27	0.41	50.15	2395
3	37.46	4.58	4.56	2694.57	0.45	798.39	0.38	50.10	598307
Year = 1999									
1	37.87	4.48	4.45	2693.86	0.51	454.98	0.39	50.18	3586
2	37.84	4.50	4.46	2670.60	0.52	410.43	0.39	50.18	3586
3	37.54	4.60	4.56	2765.73	0.45	774.33	0.38	50.15	593215
Year = 2000									
1	37.86	4.56	4.54	2386.83	0.63	115.19	0.33	50.21	2281
2	37.92	4.56	4.54	2419.19	0.60	110.06	0.35	50.21	2281
3	37.67	4.61	4.58	2815.13	0.44	769.31	0.38	50.09	595698
Year = 2001									
1	38.54	4.46	4.46	2647.18	0.61	109.16	0.41	50.33	3242
2	38.47	4.45	4.45	2678.34	0.61	109.69	0.42	50.33	3242
3	37.84	4.60	4.59	2850.53	0.43	773.61	0.39	50.13	601397
Year = 2002									
1	37.69	4.55	4.52	2746.38	0.55	463.05	0.36	50.18	3806
2	37.65	4.50	4.48	2669.42	0.50	430.96	0.38	50.18	3806
3	37.99	4.60	4.58	2865.74	0.42	760.07	0.39	50.15	610195
Year = 2003									
1	38.14	4.54	4.53	2888.71	0.66	302.39	0.41	50.19	3109
2	38.10	4.54	4.53	2885.42	0.63	258.92	0.42	50.19	3109
3	38.07	4.60	4.58	2854.84	0.42	744.26	0.40	50.12	618765
Year = 2004									

(continued)

group	Age	$\log(w_{t-1})$	$\log(w_{t-2})$	Job Tenure	Manufact.	Firm Size	Female	propensity	N
1	37.42	4.55	4.53	2282.32	0.34	195.11	0.41	50.13	2062
2	37.40	4.55	4.52	2292.93	0.33	195.41	0.41	50.13	2062
3	38.18	4.61	4.59	2888.77	0.43	745.39	0.39	50.08	606139
Year = 2005									
1	38.35	4.47	4.47	2625.83	0.58	80.69	0.36	50.21	1674
2	38.58	4.46	4.46	2522.90	0.55	78.47	0.37	50.21	1674
3	38.28	4.60	4.60	2871.40	0.42	774.47	0.39	50.07	611101
Year = 2006									
1	38.59	4.53	4.50	2699.59	0.44	146.52	0.38	50.11	1681
2	38.31	4.55	4.52	2662.30	0.44	143.97	0.38	50.11	1681
3	38.38	4.60	4.57	2799.29	0.40	813.65	0.40	50.06	643058
Year = 2007									
1	37.30	4.30	4.28	2346.15	0.38	121.21	0.47	50.13	1250
2	36.86	4.33	4.30	2327.19	0.35	113.18	0.47	50.13	1250
3	38.50	4.62	4.58	2791.22	0.40	819.37	0.40	50.05	661829
Year = 2008									
1	37.53	4.44	4.43	2091.03	0.46	261.22	0.41	50.15	2564
2	37.74	4.45	4.44	2132.42	0.47	249.49	0.41	50.15	2564
3	38.65	4.61	4.60	2797.96	0.40	823.29	0.40	50.09	673896
Year = 2009									
1	38.70	4.51	4.48	2582.17	0.65	168.67	0.32	50.33	4127
2	39.01	4.52	4.48	2624.11	0.65	164.29	0.33	50.33	4127
3	38.71	4.62	4.59	2794.65	0.40	792.83	0.40	50.15	686895

B. The Average Cost of Job Displacement

The main outcome variables we consider are total annual earnings, total annual days employed, and log daily wage from the employer on January 1st each year. Figure 15 shows the estimated causal effect of job-displacement for these three outcome variables using the event study estimation equation 2. In the year after job displacement, earnings losses amount to approximately €12,000, or 26 percent of pre-displacement earnings. In the following years earnings increase, but the recovery fades out after 5-6 year, after which the losses still amount to over €5,000 yearly, or 13% in terms of pre-displacement earnings. Figure 15 further shows that this decline in earnings both stem from employment losses and declines in log-wages. In the two years after job displacement, employment losses amount to almost 60-80 days. After a quick recovery, employment losses stabilize around 20 days per year. More strikingly, there is essentially no recovery in log-wages. Displaced workers' wages decline by 12%, with very little recovery in the first 10 years after job-displacement. The evolution of earnings losses looks surprisingly similar to those in the US (see *e.g.*, Davis and Von Wachter, 2011), despite the institutional differences between the US and Austria.

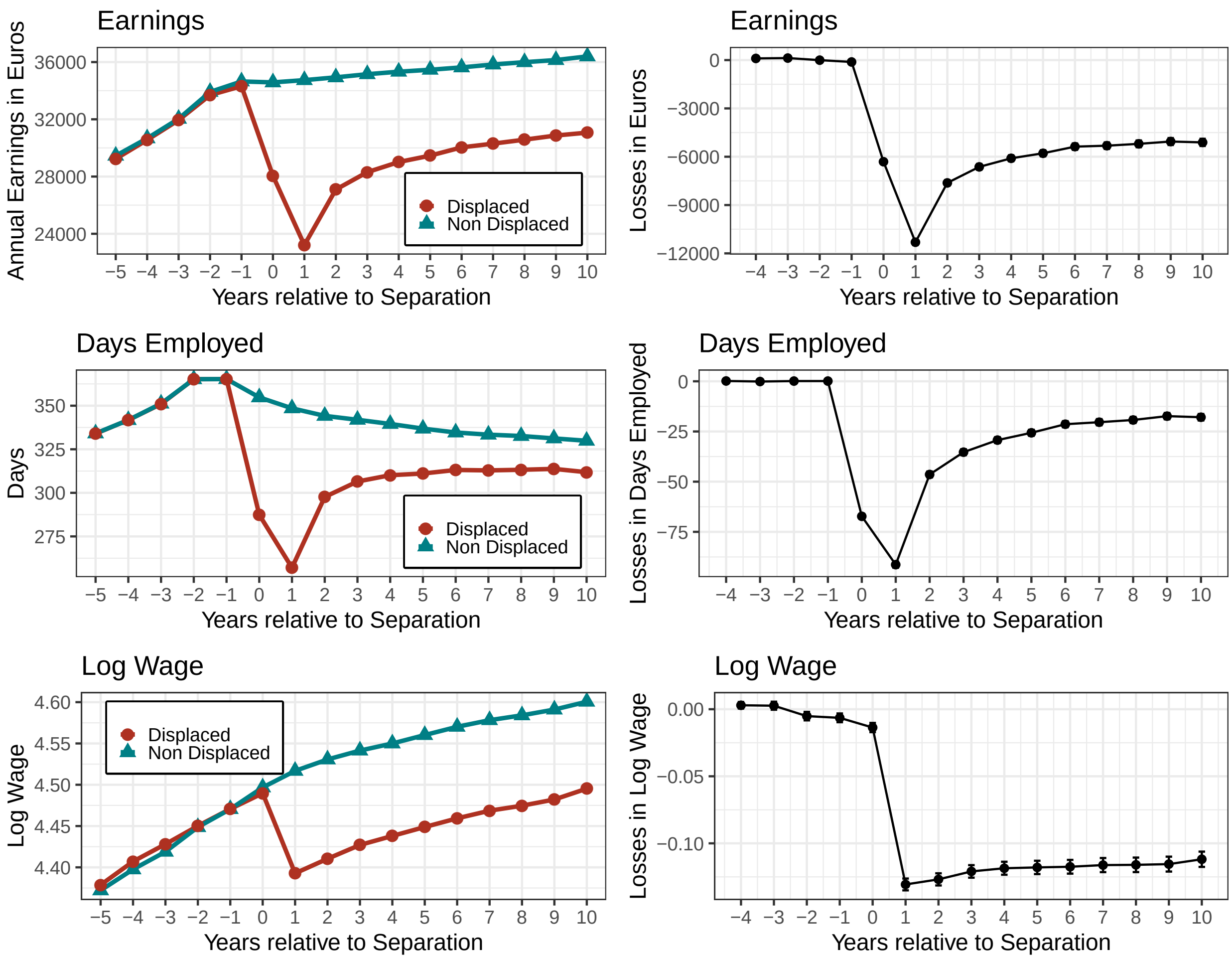


Figure 15: Earnings Losses of displaced workers - Eventstudy regression estimates of equation (2). Period 0 corresponds to the separation year. Earnings and days employed are computed for the whole year, log-wages are computed as the log average daily wage from the employer on 1st January. Control group is selected via propensity score matching.

C. Data Appendix

We use the labor market data base provided by the Austrian social security agency. The data comprises all the relevant information to compute all benefits covered under social security in Austria. These include benefits related to old-age, unemployment, sick-leave, and maternity/paternity leave. Thus the dataset contains many overlapping spells that are not necessarily related to the labor market state of a worker. We follow the recommendations in the data manual provided by the data provider to eliminate overlapping spells and thus define unique labor market states for workers. For overlapping employment spells, we select the spell with the higher yearly income to define a unique employer at each point in time for workers.[33]

A. Computation of AKM model

We follow closely Card *et al.* (2013) for the computation of the AKM model. We use the universe of all employment spells covered by social security from 1984-2019. We select individuals aged 20-60. For each year and individual, we select the worker-establishment pair with the highest income in a given year. This yields 108,050,149 worker-year observations. We drop all observations of individuals in our earnings loss sample. This includes both displaced workers, as well as the selected control group. These reduces the sample by about 4 percent. We further extract the largest connected set, which leaves 103,872,312 worker-year observations for our estimation sample. As the left hand side variable we use the log daily wage, computed by dividing the yearly income from the dominant establishment by the days employed at that establishment. Wages are further deflated by CPI to 2017 levels, and winsorized at the 0.5 and 99.5 percentile. In addition to worker and firm fixed effects, we also control for year fixed effects and a cubic in age (the linear age term is drop because of collinearity). Table 6 presents the log-wage variance decomposition based on the AKM regression results.

B. Partitioning Variables

Table 7 summarizes all partitioning variables and their definitions. Table 8 presents the cut points used for the categorization of continuous variables and Figure 16 shows how these variables are correlated with each other in the estimation sample.

[33]Working at multiple employers is very uncommon in Austria, this applies to less then 0.1 percent of spells.

Table 6: AKM Decomposition

Person and establishment parameters		
Number of person effects	7397638	
Number of establishment effects	946155	
Variance Decomposition	Total	Share
Var(Person effects)	0.128	46.8
Var(Establ effects)	0.061	23.7
Var(Xb)	0.024	9.3
Var(Residual)	0.050	19.3
2Cov(Person/establ. effects)	0.030	11.7
2Cov(perons/Xb)	-0.022	-8.6
2Cov(establ/Xb)	-0.006	-2.2
Var(log-wages)	0.258	100
Summary of Estimation		
Adj R^2	0.790	
Sample size	103,872,312	

Table 7: Definitions of the Partitioning Variables

Variable	Definition
Pers. FE	Worker Effect from Regression equation (4) binned to deciles
Firm FE	Firm Effect from Regression equation (4) binned to deciles
Match Effect	Residual from Regression equation (4) binned to deciles
Avg. Region Firm FE	Average percentile of the firm effect from regression equation (4), leaving out the previous employer. Binned into deciles.
Austrian	Indicator for Austrian citizenship
Female	Indicator for worker's gender
Blue Collar	Indicator for blue collar employment relationship
Worker's age	Worker's age in years
Firm age	Firm's age binned to deciles
Job Tenure	Job tenure measured at the beginning of the event year, binned into deciles
Manufacturing	manufacturing as nace-1 industry
U-rate Region	Average unemployment rate 1984-2019 in NUTS-3 region of previous employer
U-Rate Ind.	Average unemployment rate 1984-2019 in nace-1 industry of previous employer
Firm Size	Number of employees on 1st of January of the event year
Market Share	Employment share of previous employer in NUTS-3 region and nace-1 industry
Herfindahl	Herfindahl-Hirschman index of labor market concentration in NUTS-3 and nace-1 industry.
Firm Sep. Rate	Average firm separation rate in the 5 years leading up to the event year, excluding recalls.
Recession	Unemployment rate above it's trend in the year of the mass layoff event
# Employers	Number of Employers before event year, observations binned above 4

Table 8

	P10	P20	P30	P40	P50	P60	P70	P80	P90
Firm Size	4	10	21	43	90	178	349	751	2210
Job Tenure	61	209	366	641	974	1440	2022	2922	4491
Herfindahl Index * 100	0.012	0.028	0.038	0.052	0.065	0.088	0.157	0.183	0.321
Market Share * 100	0.000	0.002	0.005	0.01	0.02	0.04	0.08	0.17	0.39
Avg. Firm FE	40.98	45.17	49.18	50.21	51.43	52.14	52.78	53.98	55.80
Match Effect	-0.090	-0.046	-0.026	-0.013	0.002	0.011	0.029	0.059	0.121
Industry U-Rate	0.034	0.070	0.073	0.074	0.074	0.090	0.090	0.936	0.225
Regional U-Rate	0.071	0.078	0.084	0.094	0.107	0.107	0.107	0.110	0.132
Firm Age	4	7	11	14	17	21	25	30	37
Firm Sep. Rate	0.004	0.018	0.029	0.042	0.059	0.081	0.111	0.162	0.277

Notes: Table shows the 10th to 90th percentile of the continuous variables based on the distribution of all employees on the reference day. These are used as the cut points for the categorization of continuous variables to deciles.

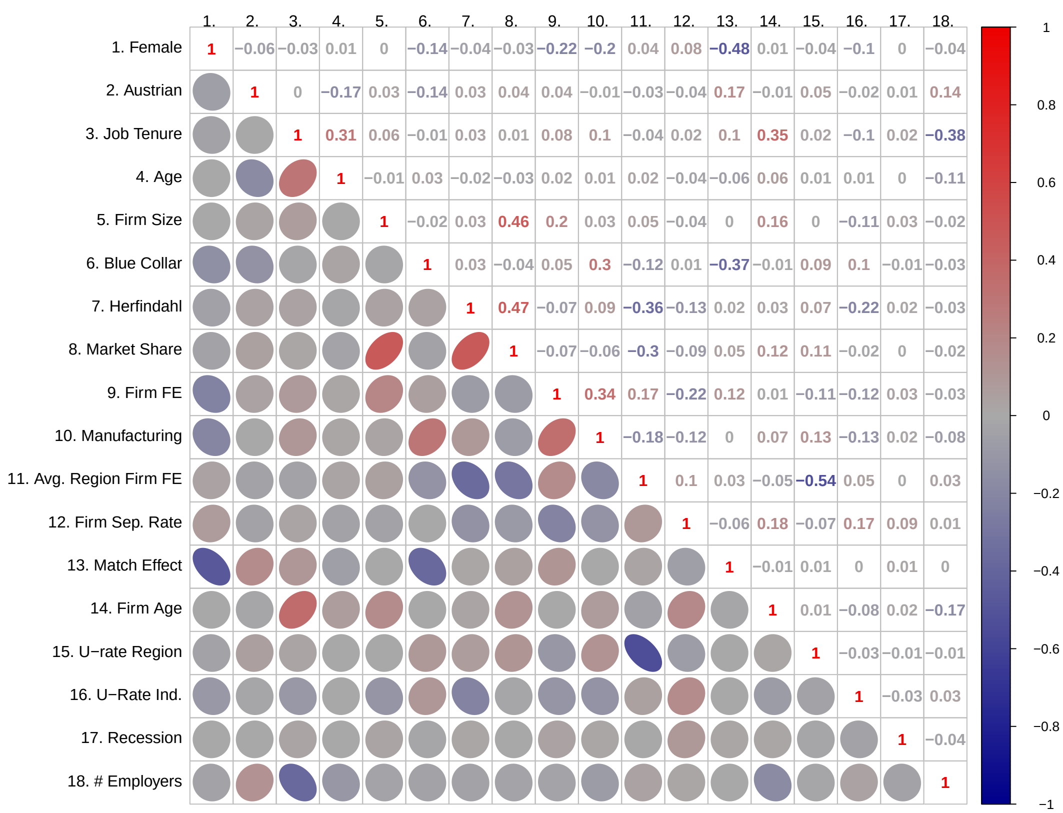


Figure 16: Correlogram of partitioning variables

D. Construction of Weights and Numerical Implementation

With the forest at hand, we can proceed with the construction of weights. Suppose that there is a forest with B trees indexed by b. Then weight $\alpha_{it}^{b}(\mathbf{z})$ measures the similarity of observation (i,t) with $\mathbf{z}$ and is defined as:

$$\alpha_{it}^{b}(\mathbf{z}) := \begin{cases} \frac{1}{|L_b(\mathbf{z})|}, & \mathbf{z}_{it} \in L_b(\mathbf{z}) \\ 0, & \text{otherwise}, \end{cases} \tag{11}$$

where $L_b(\mathbf{z})$ is the set of all observations, which share the same terminal node ("leaf") with an individual with characteristics $\mathbf{z}$ in tree b and $|L_b(\mathbf{z})|$ is the size of this set. The weight $\alpha_i(\mathbf{z})$ used in (6) is the average across all trees: $\alpha_i(\mathbf{z}) := \frac{1}{B}\sum_{b=1}^{B} \alpha_{it}^{b}(\mathbf{z})$.

As mentioned before, the forest is built to maximize the heterogeneity of treatment effects (with an additional adjustment for balanced subsamples) across splits and this is expressed by (7). That said, because of computational complexity, this criterion is replaced by more numerically efficient approximation in the spirit of gradient boosting due to Friedman (2001). However, before presenting the exact procedure, one remark should be made. In a given data partition $\mathcal{P}$, the OLS estimator trained on $\mathcal{P}$ meets the following condition:

$$\frac{1}{N_{\mathcal{P}}} \sum_{(i,t)\in\mathcal{P}} \mathbf{x}_{it}' u_{it} = \mathbf{0}_{18}. \tag{12}$$

Then the treatment effect τ_{C_k} of any subset $C_k \in \mathcal{P}$ can be approximated by:

$$\tau_{C_k} \approx \tau_{\mathcal{P}} + \xi' \left(\frac{1}{N_{\mathcal{P}}} \sum_{(j,s)\in\mathcal{P}} \mathbf{x}_{js}\mathbf{x}_{js}' \right)^{-1} \cdot \frac{1}{N_{C_k}} \sum_{(i,t)\in C_k} \mathbf{x}_{it} u_{it}, \tag{13}$$

where $\xi' = (1, \mathbf{0}_{17})$ is a vector selecting τ from the vector of all regression coefficients, and u_{it} is the residual term from the model estimated on $\mathcal{P}$.[34]

Then, the impact of an individual observation (i,t) on τ_{C_k} is given by:

$$\rho_{it} = \xi' \left(\frac{1}{N_{\mathcal{P}}} \sum_{(j,s)\in\mathcal{P}} \mathbf{x}_{js}\mathbf{x}_{js}' \right)^{-1} \cdot \mathbf{x}_{it} u_{it}. \tag{14}$$

[34] Notice that this approximation can be interpreted as an improved guess $x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}$ where the function $f()$ is given by (5), x_n is $\tau_{\mathcal{P}}$ and x_{n+1} corresponds to τ_{C_k} in the textbook Newton-Raphson root-finding algorithm.

Using the CART algorithm by Breiman *et al.* (1984) on transformed outcomes (14), we are able to find such a split into C_1 and C_2 which minimizes the within-group sum of squares of ρ. Using the fact that the grand mean of ρ in the parent node is equal to zero, this implies that the algorithm maximizes the between-group sum of squares, *i.e.*:

$$\frac{1}{N_{C_1}}\left(\sum_{(i,t)\in C_1}\rho_{it}\right)^2+\frac{1}{N_{C_2}}\left(\sum_{(i,t)\in C_2}\rho_{it}\right)^2, \tag{15}$$

which, as Athey *et al.* (2019) show for a more general case, is consistent with maximizing criterion (7). Thanks to this relabelling strategy, the whole procedure of building a forest gains substantial computational performance.

The formula for the variance of estimates can be derived, as in the standard GMM, by applying the delta method to the moment conditions of a weighted least squared regression, $f(\mathbf{z}) := \sum_{it}\alpha_{it}(\mathbf{z})\mathbf{x}_{it}\varepsilon_{it}(\mathbf{z})$. In our case we are interested only in $\hat{\tau}(\mathbf{z})$, so the whole formula is multiplied by ξ, which picks the estimate of our interest. As a result, the variance of $\hat{\tau}(\mathbf{z})$ is given by:

$$\mathrm{Var}(\hat{\tau}(\mathbf{z})) = \xi' V(\mathbf{z})^{-1}H(\mathbf{z})\left(V(\mathbf{z})^{-1}\right)'\xi. \tag{16}$$

where $H(\mathbf{z}) := f(\mathbf{z})f(\mathbf{z})'$ is the variance of $f(\mathbf{z})$ and $V(\mathbf{z}) := \nabla_{\{\tau,\theta,\gamma\}}f(\mathbf{z}) = -\sum_{it}\alpha_{it}(\mathbf{z})\mathbf{x}_{it}\mathbf{x}_{it}'$ is the Jacobian of $f(\mathbf{z})$. That said, by no means Equation (16) should be estimated by simply using training observations, just like in the traditional GMM. The underlying reason for that is this would ignore the whole model selection step, which give rise to values of $\alpha_{it}(\mathbf{z})$ and $\varepsilon_{it}(\mathbf{z})$ in the presented formula. To circumvent this concern, as suggested by Athey *et al.* (2019), we employ a so-called bootstrap of little bags in the spirit of Sexton and Laake (2009) to evaluate $H(\mathbf{z})$. This procedure involves computing a between-group variance of $\hat{\tau}(\mathbf{z})$, where trees are pooled into bags and built using the same bootstrap subsample. Using one-way ANOVA it can be shown that this measure is approximately equal to (16). Thanks to this, our standard errors measure estimation accuracy affected by both machine-learning uncertainty and estimation noise.

E. Decomposing Earnings Losses into Employment and Wage Losses

Losses in earnings are a resultant of losses in their two margins, wages and employment. For this reason, we quantify which fraction of earnings losses originate from declines in employment and wage losses, and how this decomposition differs across workers.

Formally, we can write annual earnings as the product of the number of days employed multiplied by the daily wage in a that year, *i.e.* $y = N_d w$. We follow Schmieder *et al.* (2023), and decompose earnings losses into losses stemming from working fewer days and losses in daily wages. The wage gap between displaced workers and their control group $\Delta = y^C - y^D$ can be decomposed into three terms the following way:

$$
\begin{aligned}
\mathbb{E}[\Delta] &= \mathbb{E}[y^C] - \mathbb{E}[y^D] = \mathbb{E}[N_d^C w^C] - \mathbb{E}[N_d^D w^D] \\
&= \mathbb{E}[N_d^C]\mathbb{E}[w^C] - \mathbb{E}[N_d^D]\mathbb{E}[w^D] + \mathrm{Cov}(N_d^C, w^C) - \mathrm{Cov}(N_d^D, w^D) \\
&= \left(\mathbb{E}[N_d^C] - \mathbb{E}[N_d^D]\right)\mathbb{E}[w^C] + \mathbb{E}[N_d^D]\left(\mathbb{E}[w^C] - \mathbb{E}[w^D]\right) + \Delta\mathrm{Cov}(N_d, w) \\
&= \Delta\mathbb{E}[N_d]\mathbb{E}[w^C] + \mathbb{E}[N_d^D]\Delta\mathbb{E}[w] + \Delta\mathrm{Cov}(N_d, w).
\end{aligned} \tag{17}
$$

Figure 17 shows this decomposition by quartile of predicted treatment effect. Overall, losses in days employed contribute a significant fraction in the short run to overall earnings losses, but the long run persistent losses are almost entirely driven by changes in wages. In the first year after separation, employment losses explain approximately one third of earnings losses. But the contribution of employment losses fades away quickly over time as workers transition back to work. In the long run, losses in days employed only contribute around 10%, and earnings losses are almost entirely driven by losses in wages. The change in the covariance term counteracts earnings losses. This implies, that shortly after displacement, workers with higher wages are employed more days per week compared to the control group. The decomposition results are very similar across the different treatment effect groups except the group with the lowest earnings losses. For this group, short term losses are entirely driven by employment losses, whereas they even experience wage gains in the long run.

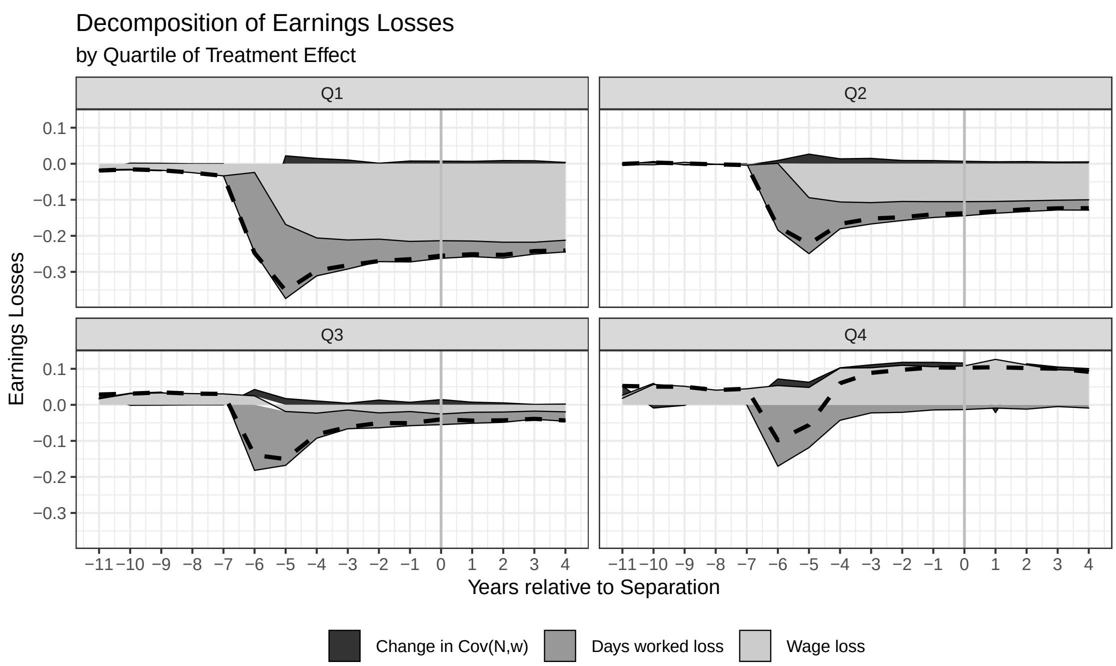
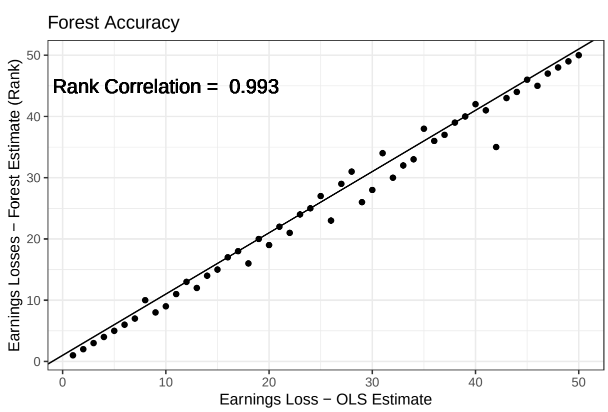


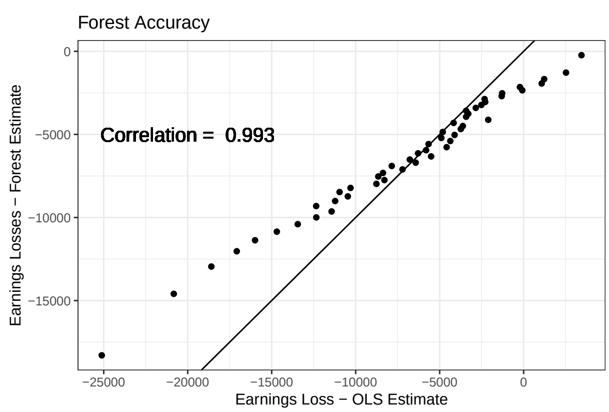
Figure 17: Decomposition of earnings losses by quartile of predicted treatment effect using equation (17). Estimates from a generalized random forest. Broken line indicates total earnings losses.

F. Accuracy of the Random Forest

How accurate are the estimates of the random forest? Evaluating the accuracy is not a straightforward task. We are not estimating an observed outcome, but a treatment effect. Thus, there is no ground truth which we can use to evaluate the estimates. To nevertheless provide a measure of accuracy, we execute the following exercise. First, using the estimates by our random forest, we bin individuals into 50 groups based on their estimated earnings losses. For each of these groups, we separately estimate equation (2) using OLS. We then compare the OLS earnings loss estimates with the results of our random forest. The left panel of Figure 18 compares the rank correlation between the two approaches. Both, the OLS estimates and the random forest rank the groups almost in the same way, the rank correlation being 0.997. The right panel plots the OLS estimates against the earnings loss estimates from the random forest. The correlation is with 0.994 equally high. A closer inspection reveals that the earnings loss estimates from the random forest are somewhat regularized, meaning that the OLS estimates suggest a higher level of heterogeneity. We think of this as a feature, rather than a shortcoming. OLS is going to overfit towards outliers, whereas the bootstrapping estimation procedure of the random forest is only picking up heterogeneity that consistently occurs across the bootstrapped samples (bags). Put differently, the random forest only identifies predictable heterogeneity.



(a) Rank scatter plot of forest estimates against OLS outcomes



(b) Bin scatter plot of forest estimates against OLS outcomes

Figure 18: Bin scatter plot of random forest accuracy. We bin all individuals by their estimated treatment effects into 50 bins. For these 50 subgroups, we compute the OLS regression and plot the estimated cost of job displacement against the average forest estimates

G. Are the sample restrictions selecting high loss individuals?

Given the heterogeneity we document, the question arises whether the usually applied sample restrictions select individuals with particularly high earnings losses. In fact, a long standing concern in the earnings loss literature is about generalizability of the results to the whole population. With the use of our random forest we are able to address this question by predicting earnings losses for a random subset of 1 million individuals not satisfying the sample restrictions on firm size or tenure. Table 9 shows that the distributions of in-sample and out-of-sample predictions are surprisingly similar. While median losses are somewhat lower for workers not satisfying the sample restrictions, they also exhibit worker and job characteristics that lead to more extreme earnings losses. Overall, the sample restrictions do not seem to select workers that are bound to experience significantly higher losses.

Percentile	P10	P25	P50	P75	P90
Displaced Worker Sample	-2.71	-2.11	-1.52	-1.02	-0.11
Out-of-sample Population	-3.15	-2.08	-1.35	-0.87	-0.34

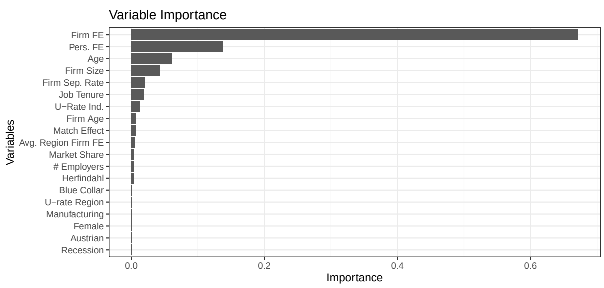
Table 9: Distribution of earnings losses relative to prior income in the displaced worker sample and in the population either not satisfying firm size or tenure restriction (1 million random subsample).

H. Variable importance using variance decomposition

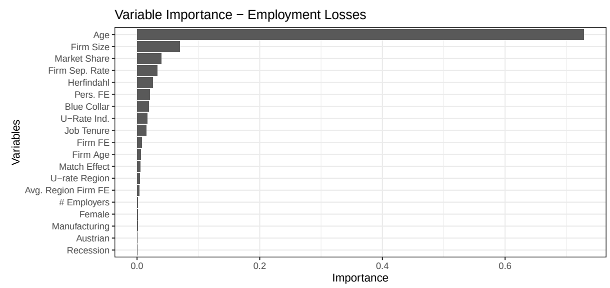
Another way to judge the statistical importance of the different channels is to estimate for how much variation in earnings losses each individual factor accounts. In order to estimate this, we project all variables onto the individual cost of job loss $\hat{\tau}_i^{\text{🌳}}$ estimated by our random forest and perform a standard variance decomposition. In practise, we estimate the following model:

$$\hat{\tau}_i^{\text{🌳}} = \alpha_0 + \sum_k \sum_j \gamma_{kj} z_i^k(j) + \epsilon_i. \tag{18}$$

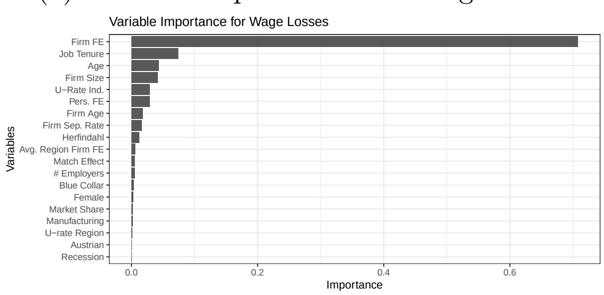
Figure 20 depicts the variance decomposition of the estimated model for earnings, employment, log-wage and firm wage premia losses. Here again, the firm wage premia have the overwhelming contribution for earnings, wage and firm wage premia losses. Its variance alone accounts for over 30% of variability in individual earnings losses and a staggering 73% of the variation in individual firm wage premia losses. Both measures of variable importance indicate that lossses in firm wage premia is the most important factor in explaining earnings, wage and firm wage losses. For employment losses on the other hand, worker's age is the most important factor determining the level of losses. Overall, the variance decomposition brings to very similar conclusions as indicated by the variable importance measure. This is quite reassuring as both metrics are constructed in a very different way. Nevertheless, as yet we are not able to say anything about the net impact of one variable on losses, in terms of magnitude and direction. To this end, we seek to pin down the counterfactual changes in the value of one variable while keeping all others at their empirical level. This is the purpose of our in-depth analysis in the next subsection.



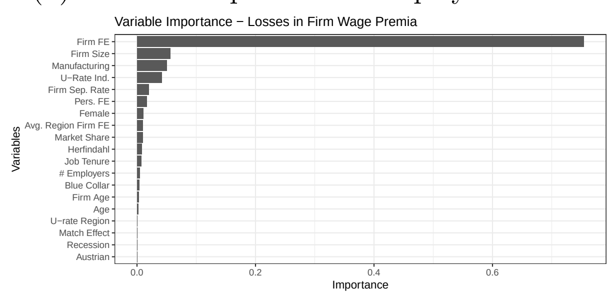
(a) Variable Importance - Earnings Losses



(b) Variable Importance - Employment Losses



(c) Variable Importance - Wage Losses



(d) Variable Importance - Firm FE Losses

Figure 19: Depth-adjusted variable frequency in splits in the GRF with a decay exponent equal to -2 and the maximum depth level of nodes equal to 4. All values sum to 1.

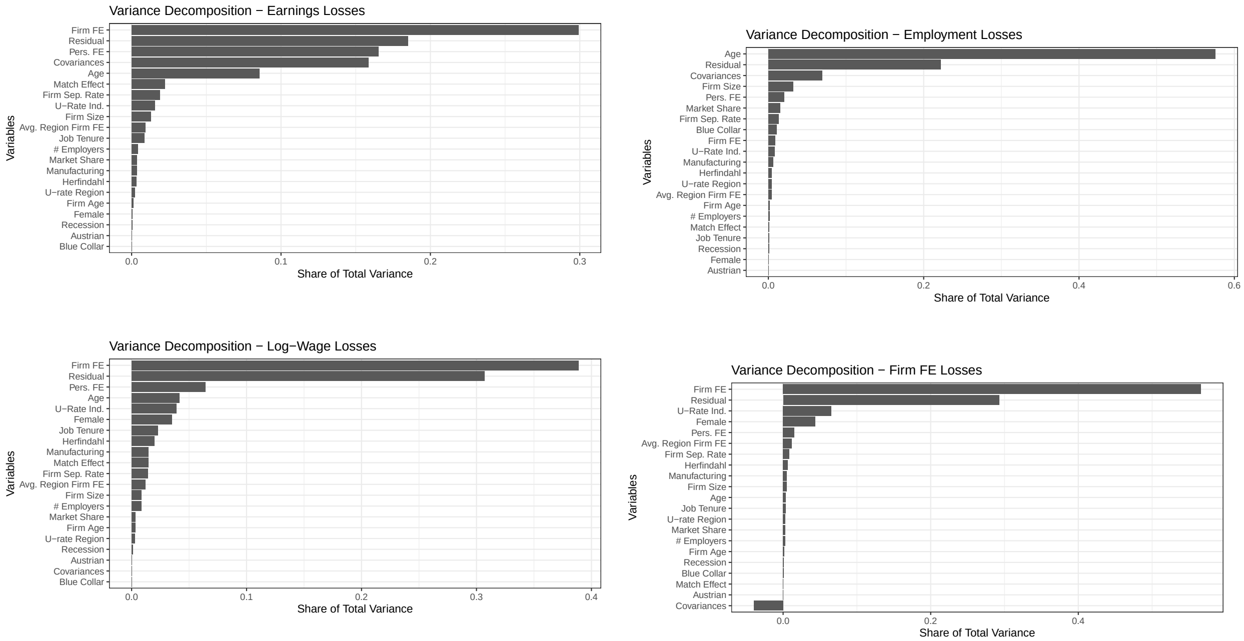


Figure 20: Variance decomposition for cost of job loss in terms of yearly earnings and log-wages. Calculations based on regression model (18)

I. Who Losses More?

Which group of workers face larger than average earnings losses, and which workers are unscarred by job displacement? To address this question, Table 10 reports descriptive statistics broken down by quartile of estimated earnings losses. The workers with the highest earnings losses have above average tenure and person effect, are employed at better paying firms, and are more likely to work in the manufacturing sector and have a white collar occupation. It is notable how different the average firm pay is across these four groups. While workers facing the highest losses work on average for firms that are paying above the eighth decile, the ones with the lowest losses are employed on average in the third firm pay decile. These two groups also differ significantly in their age. Workers with the highest losses are on average 7 years older than workers with the lowest losses. While it is interesting to understand the composition of workers with high earnings losses, these documented differences still do not address which of the factors are the driving forces behind earnings losses. Many of these variables are correlated with each other (see Figure 16), so it is hard to draw definite conclusions from these compositional differences. This question is tackled in section VI.

Table 10

	Q1	Q2	Q3	Q4
Blue Collar	0.424	0.524	0.612	0.613
Manufacturing	0.651	0.613	0.574	0.367
Austrian	0.760	0.769	0.754	0.712
Female	0.232	0.306	0.432	0.652
Firm Size	6.604	6.008	5.787	5.645
Firm Age	6.254	5.776	5.749	5.448
Firm FE	8.112	7.451	6.080	3.746
Pers. FE	7.172	6.103	4.833	3.213
Match Quality	5.963	5.358	4.982	5.087
Job Tenure	8.088	7.448	7.173	6.822
Herfindahl Index	5.048	4.728	4.712	4.564
Market Share	5.149	5.149	5.149	5.149
Avg. Firm FE	5.651	5.855	5.524	5.480
Industry UE-Rate	4.741	5.475	5.552	5.797
Regional UE-Rate	5.135	5.088	5.211	4.998
Recession	0.746	0.714	0.715	0.709
Number of Firms	2.096	2.467	2.526	2.566
Age	41.980	38.042	35.649	35.030
Firm Sepa. Rate	5.521	6.337	6.655	7.255

Notes: Table shows mean baseline characteristics for each quartile of estimated earnings losses. Predictions from a causal forest

J. Partial Effects

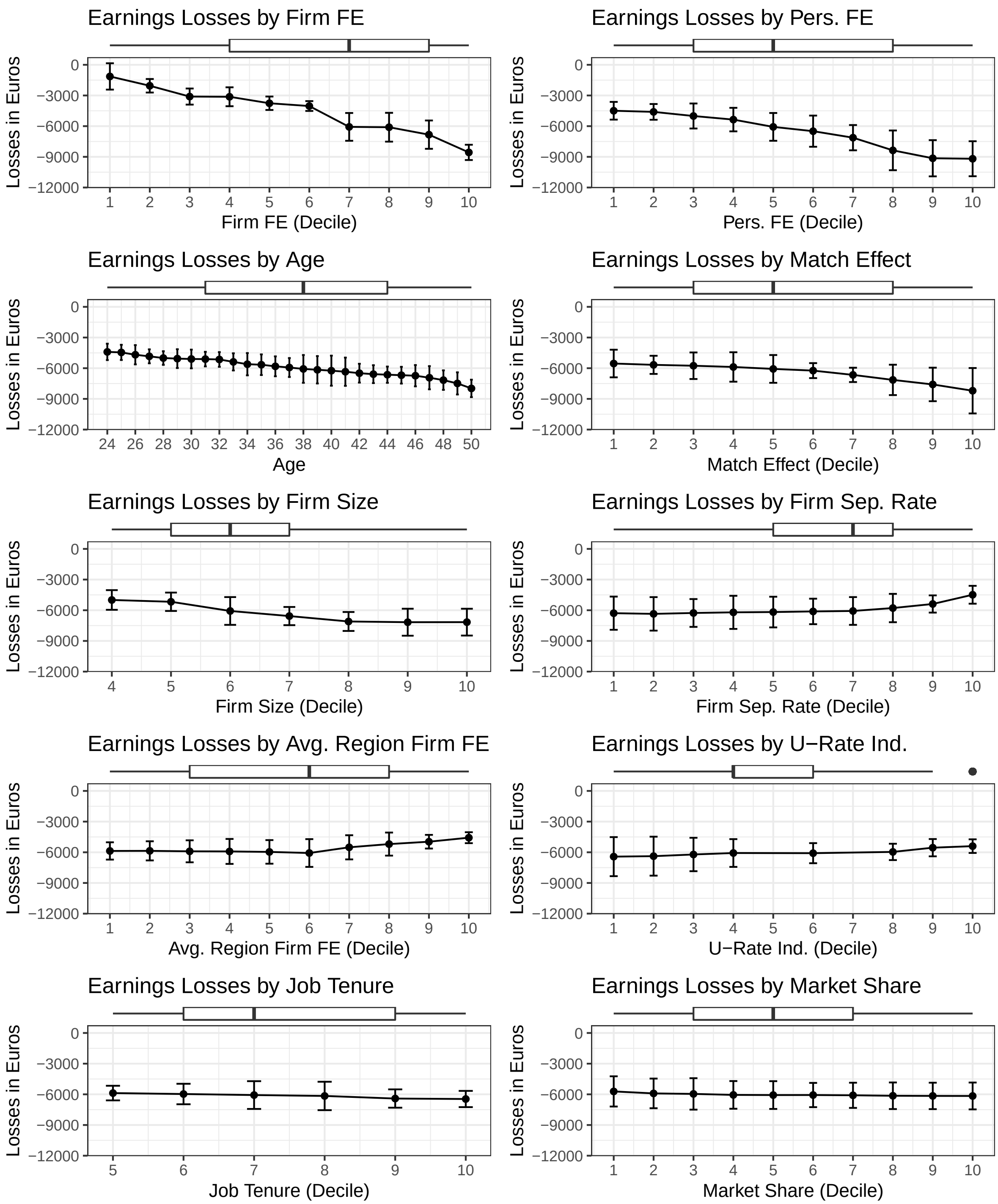


Figure 21: GRF estimates with 95% CI of losses in earnings by partitioning variables. All other variables are set to their median values. The boxplots present the distribution of the partitioning variable in the dataset (for the details see Subsection B of Section II).

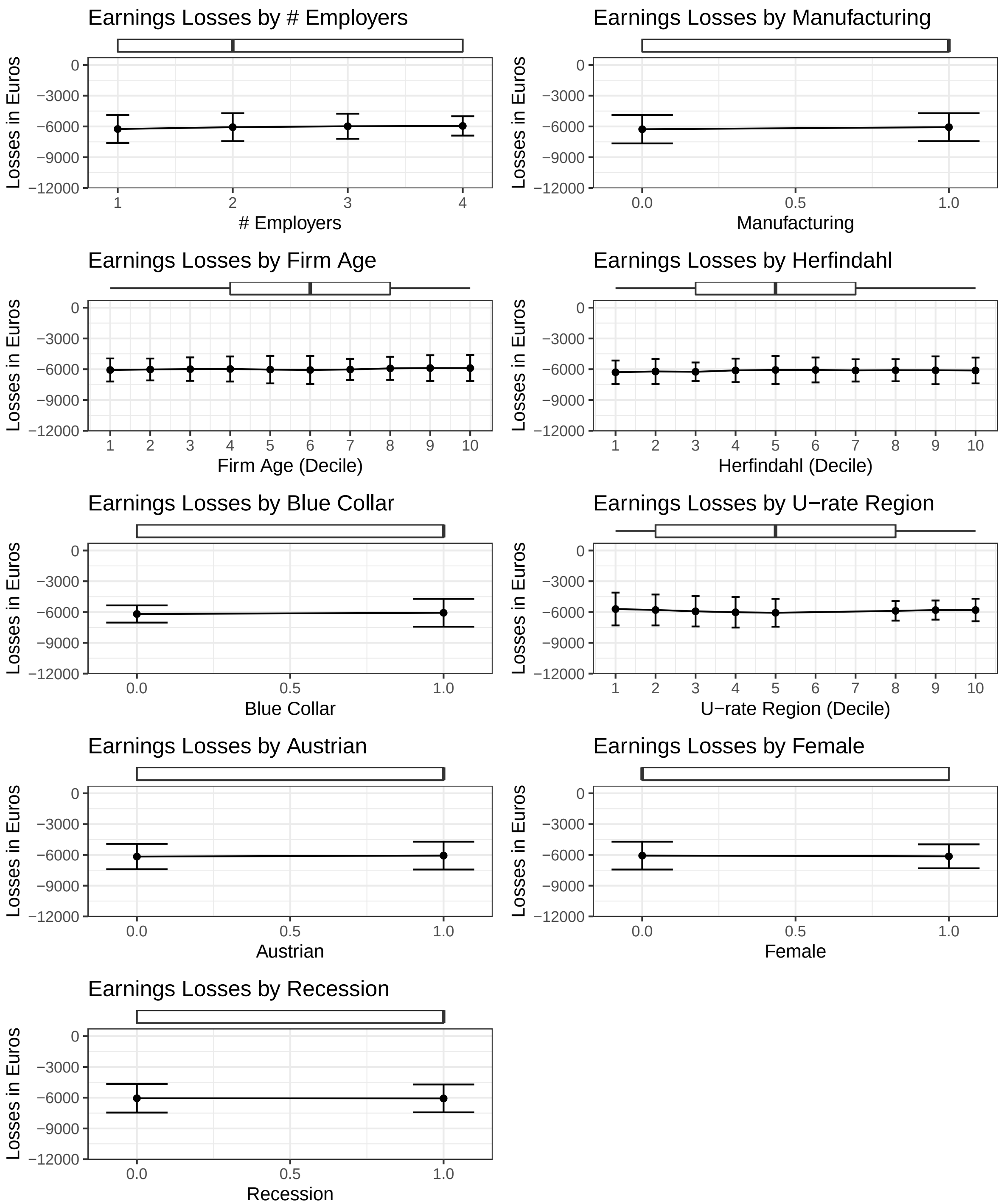


Figure 22: GRF estimates with 95% CI of losses in earnings by partitioning variables. All other variables are set to their median values. The boxplots present the distribution of the partitioning variable in the dataset (for the details see Subsection B of Section II).

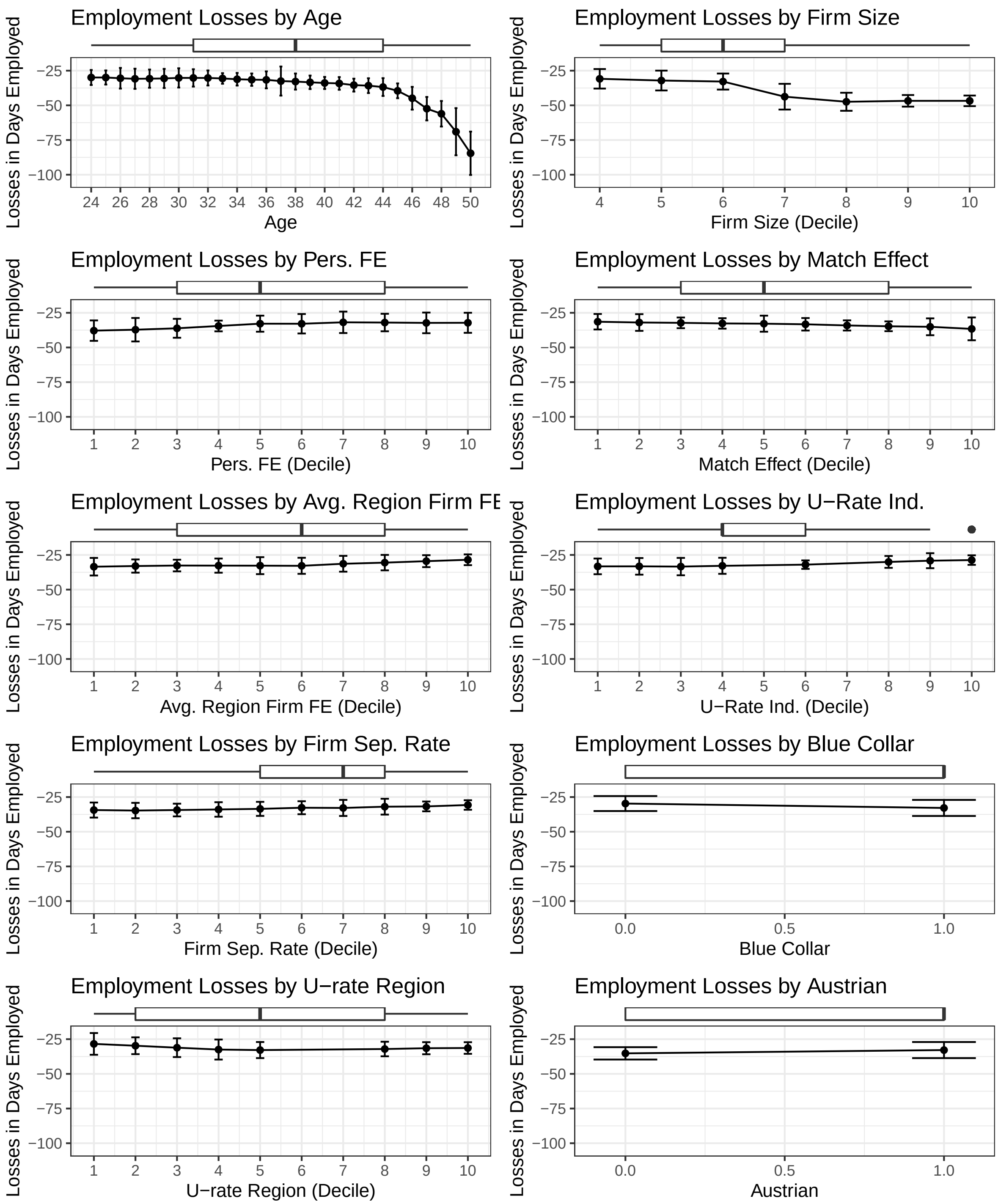


Figure 23: GRF estimates with 95% CI of employment losses by partitioning variables. All other variables are set to their median values. The boxplots present the distribution of the partitioning variable in the dataset (for the details see Subsection B of Section II).

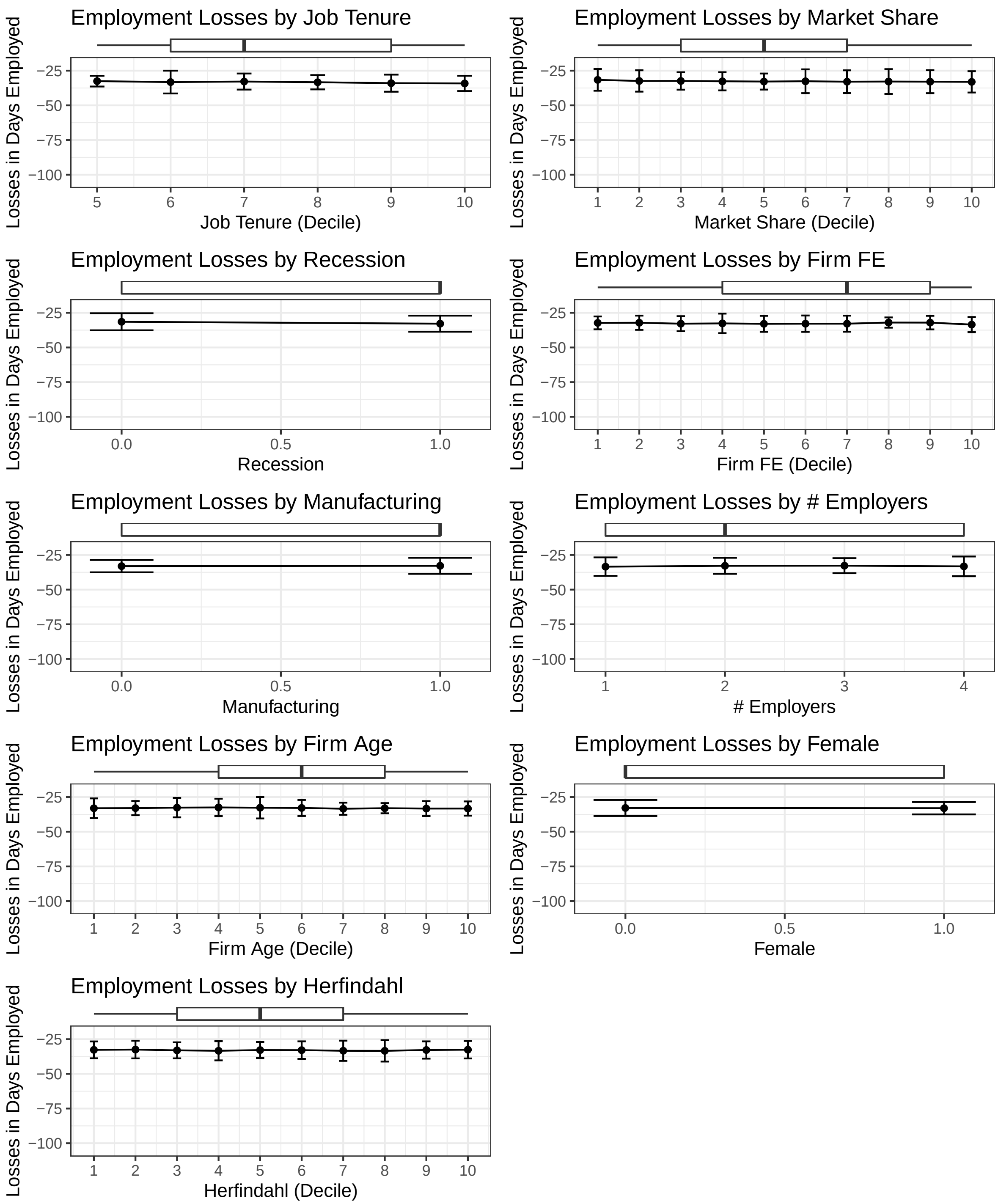


Figure 24: GRF estimates with 95% CI of employment losses by partitioning variables. All other variables are set to their median values. The boxplots present the distribution of the partitioning variable in the dataset (for the details see Subsection B of Section II).

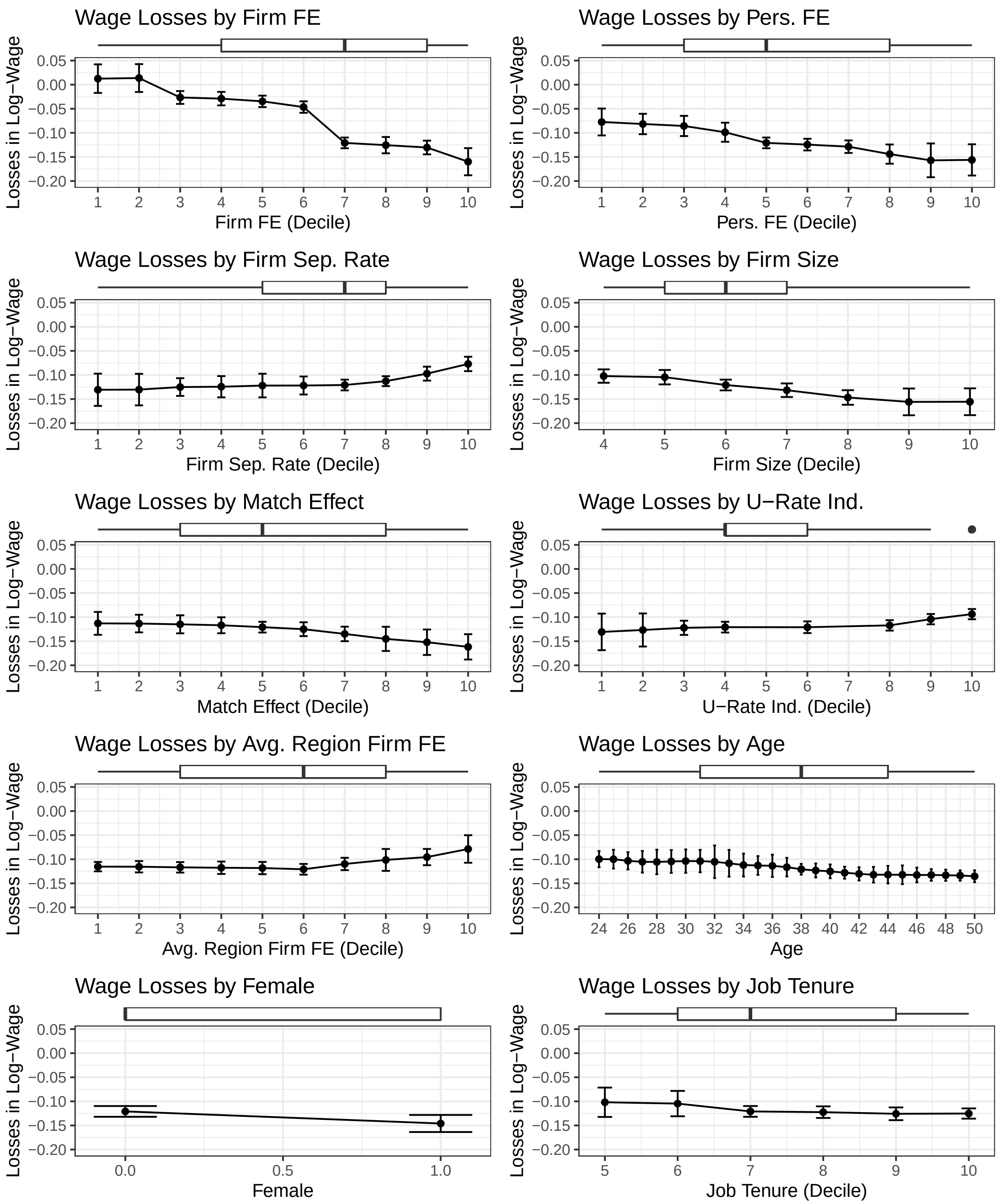


Figure 25: GRF estimates with 95% CI of wage losses by partitioning variables. All other variables are set to their median values. The boxplots present the distribution of the partitioning variable in the dataset (for the details see Subsection B of Section II).

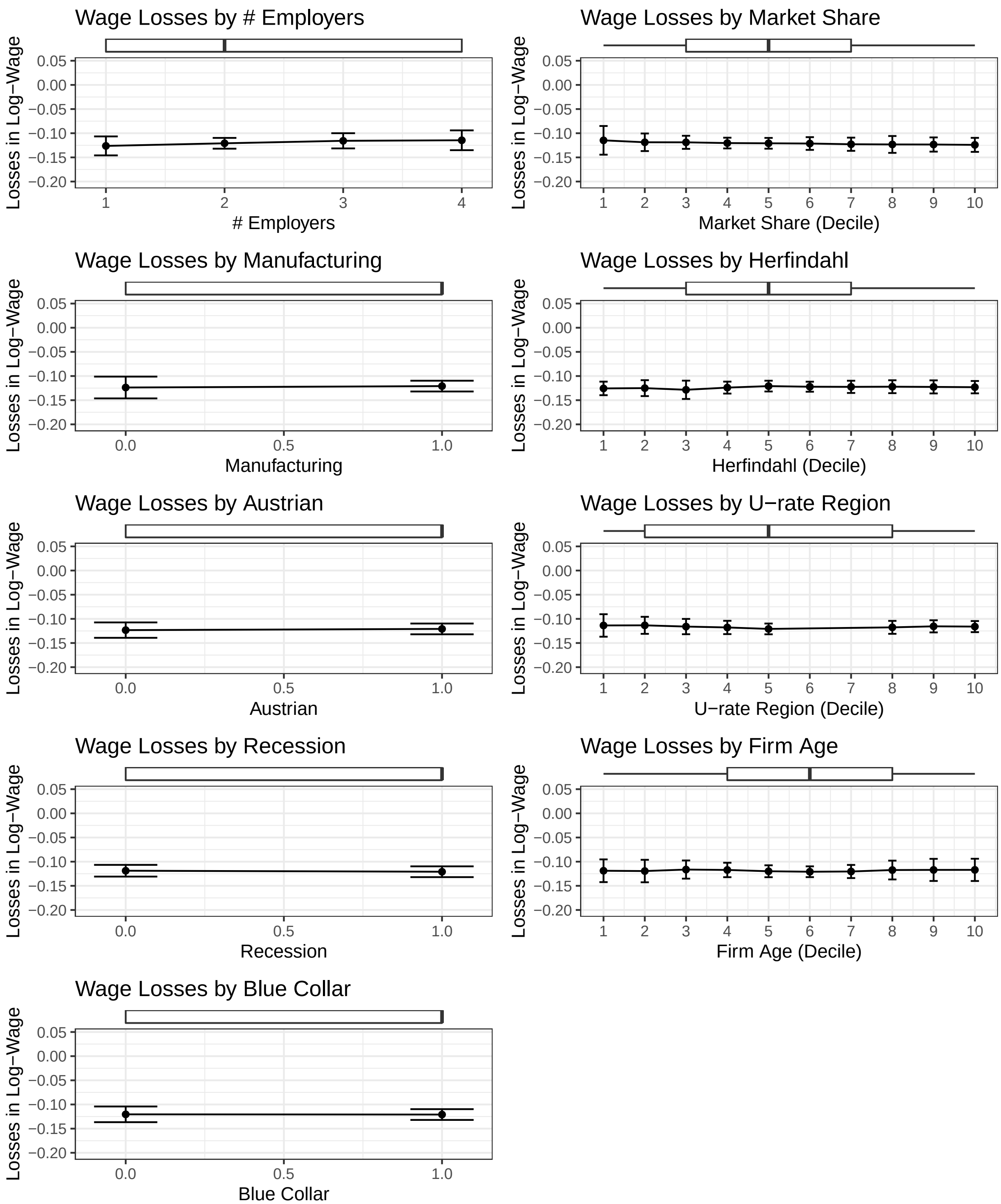


Figure 26: GRF estimates with 95% CI of wage losses by partitioning variables. All other variables are set to their median values. The boxplots present the distribution of the partitioning variable in the dataset (for the details see Subsection B of Section II).

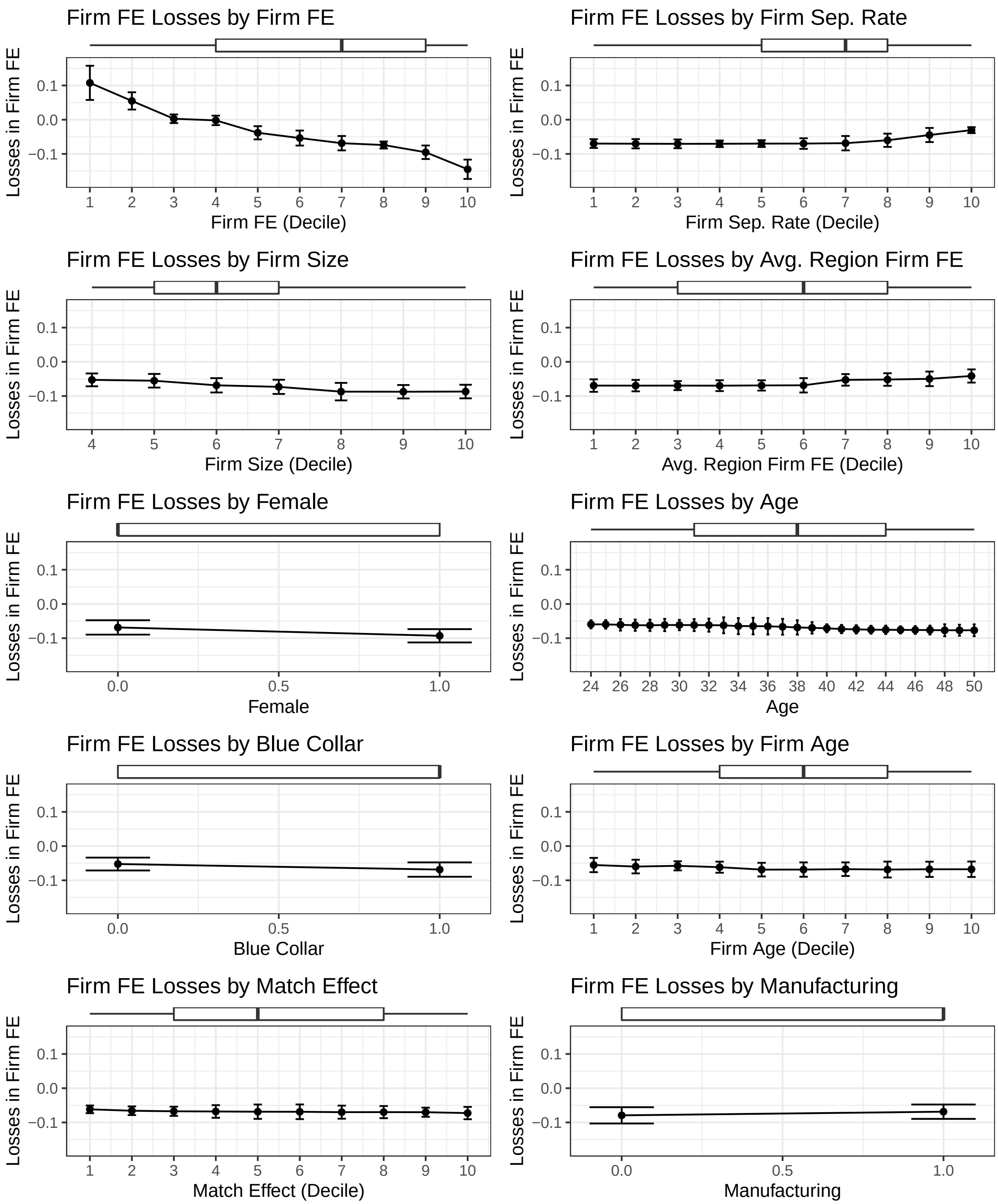


Figure 27: GRF estimates with 95% CI of firm wage premia losses by partitioning variables. All other variables are set to their median values. The boxplots present the distribution of the partitioning variable in the dataset (for the details see Subsection B of Section II).

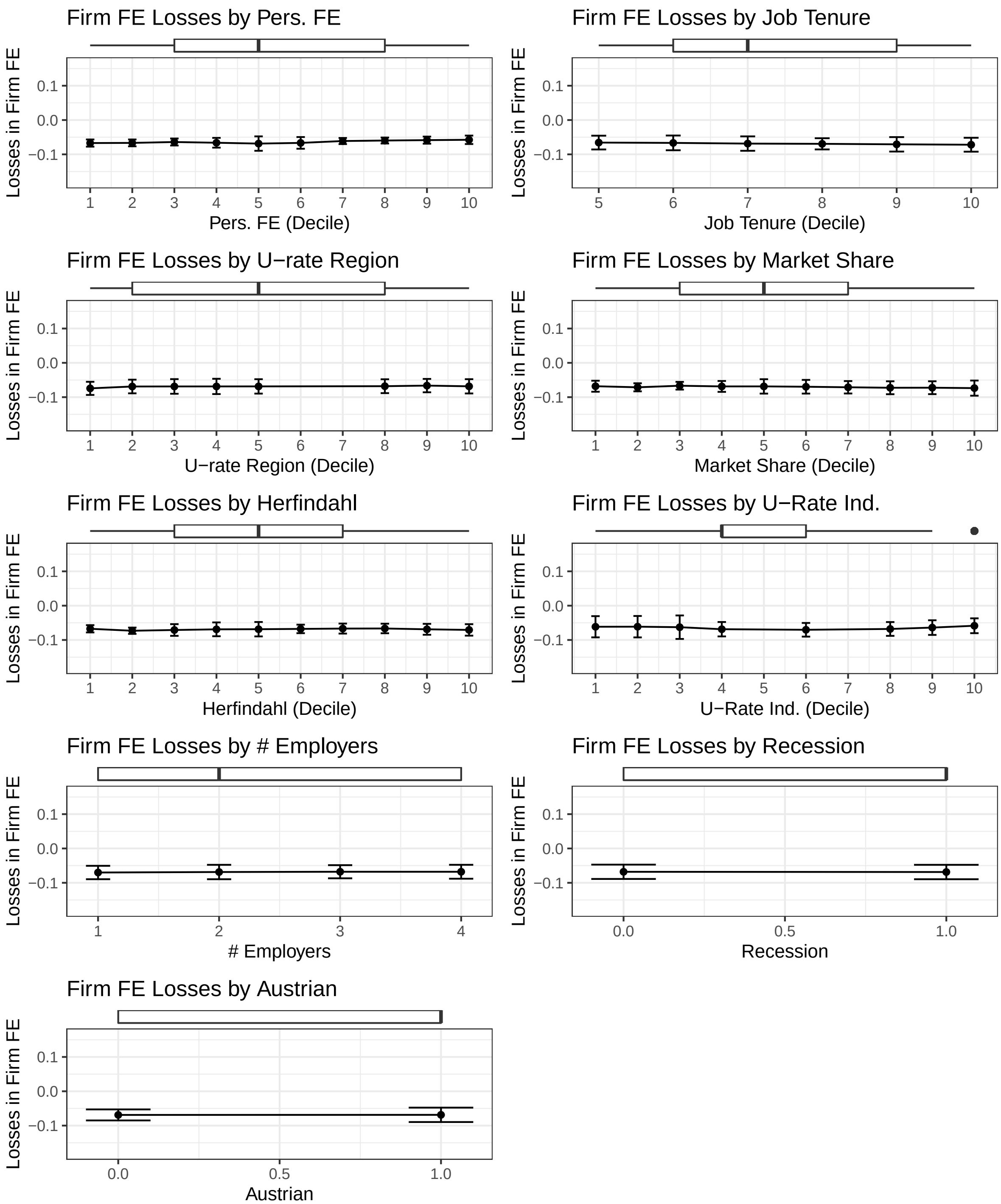


Figure 28: GRF estimates with 95% CI of losses in earnings by partitioning variables. All other variables are set to their median values. The boxplots present the distribution of the partitioning variable in the dataset (for the details see Subsection B of Section II).

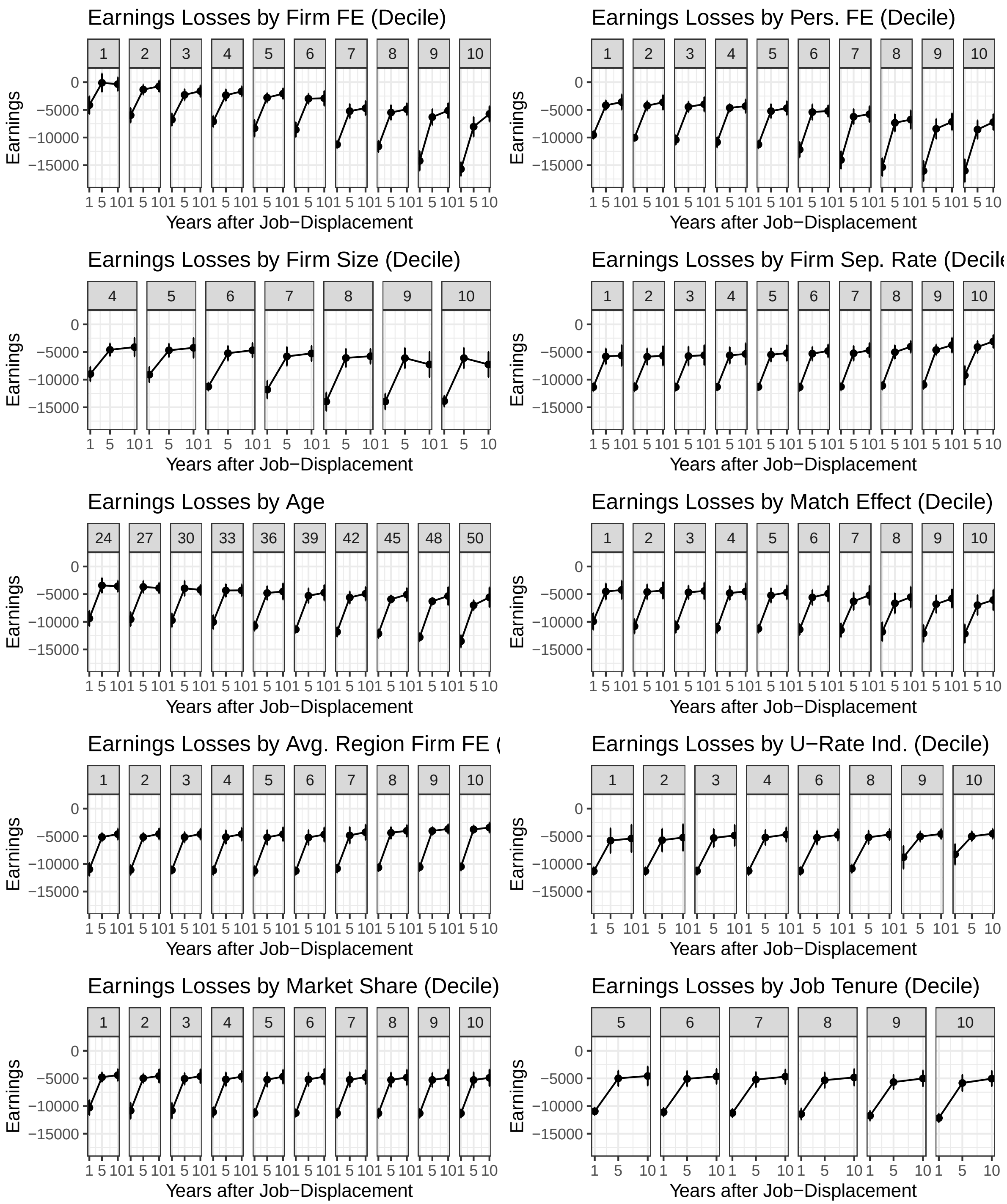


Figure 29: GRF estimates with 95% CI of losses in earnings by partitioning variables. All other variables are set to their median values. (for the details see Subsection B of Section II).

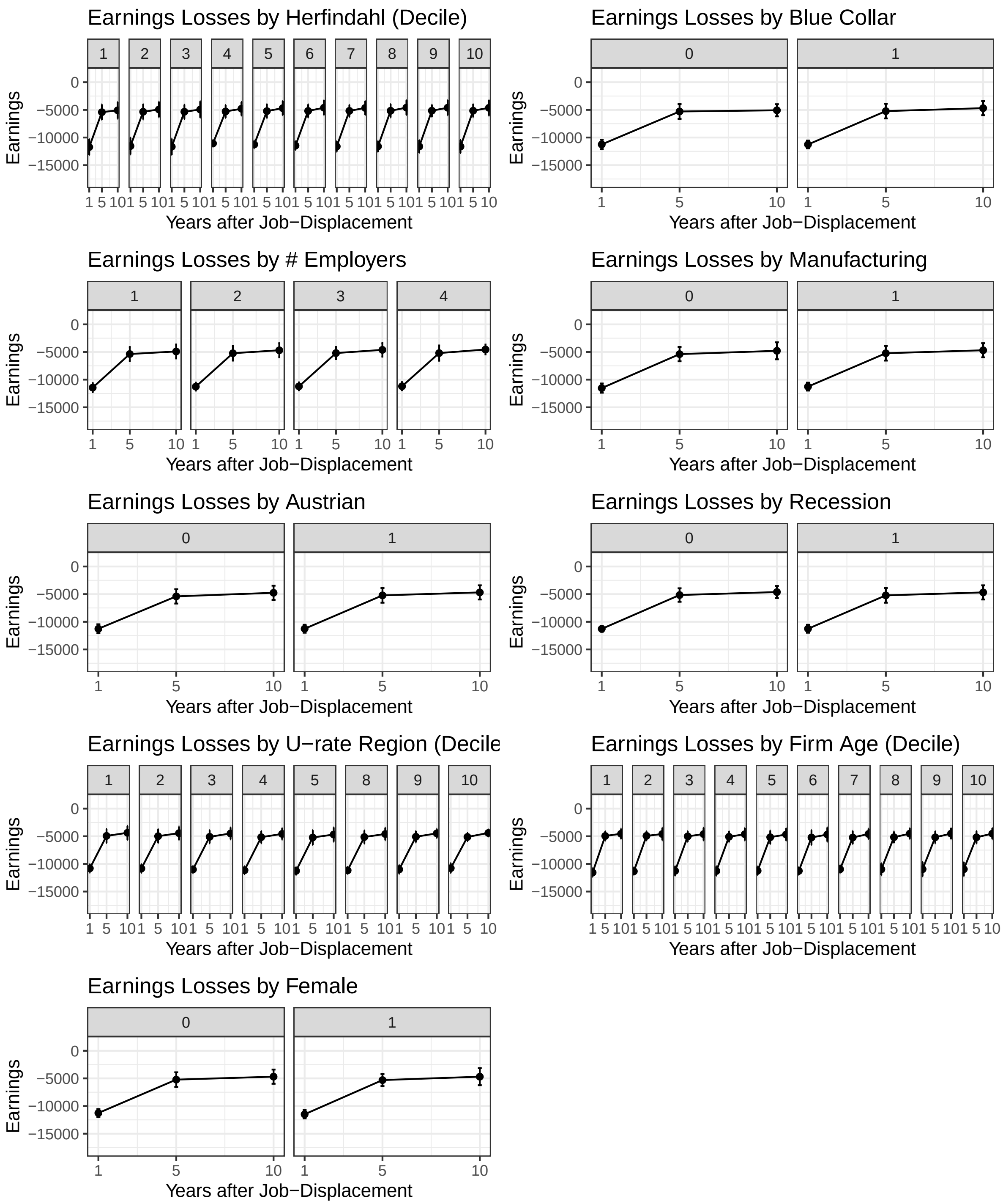


Figure 30: GRF estimates with 95% CI of losses in earnings by partitioning variables. All other variables are set to their median values. The boxplots present the distribution of the partitioning variable in the dataset (for the details see Subsection B of Section II).

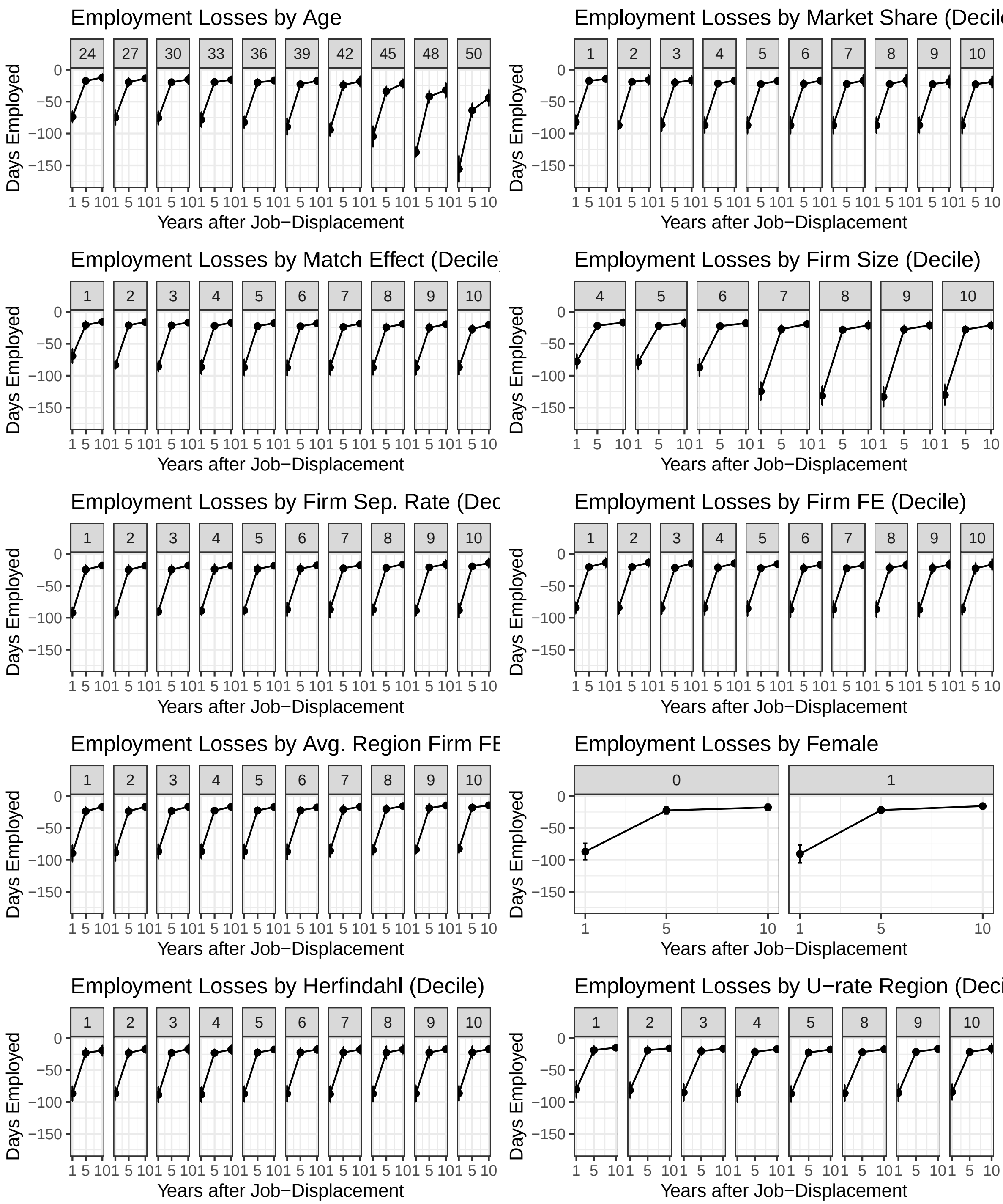


Figure 31: GRF estimates with 95% CI of employment losses by partitioning variables. All other variables are set to their median values. The boxplots present the distribution of the partitioning variable in the dataset (for the details see Subsection B of Section II).

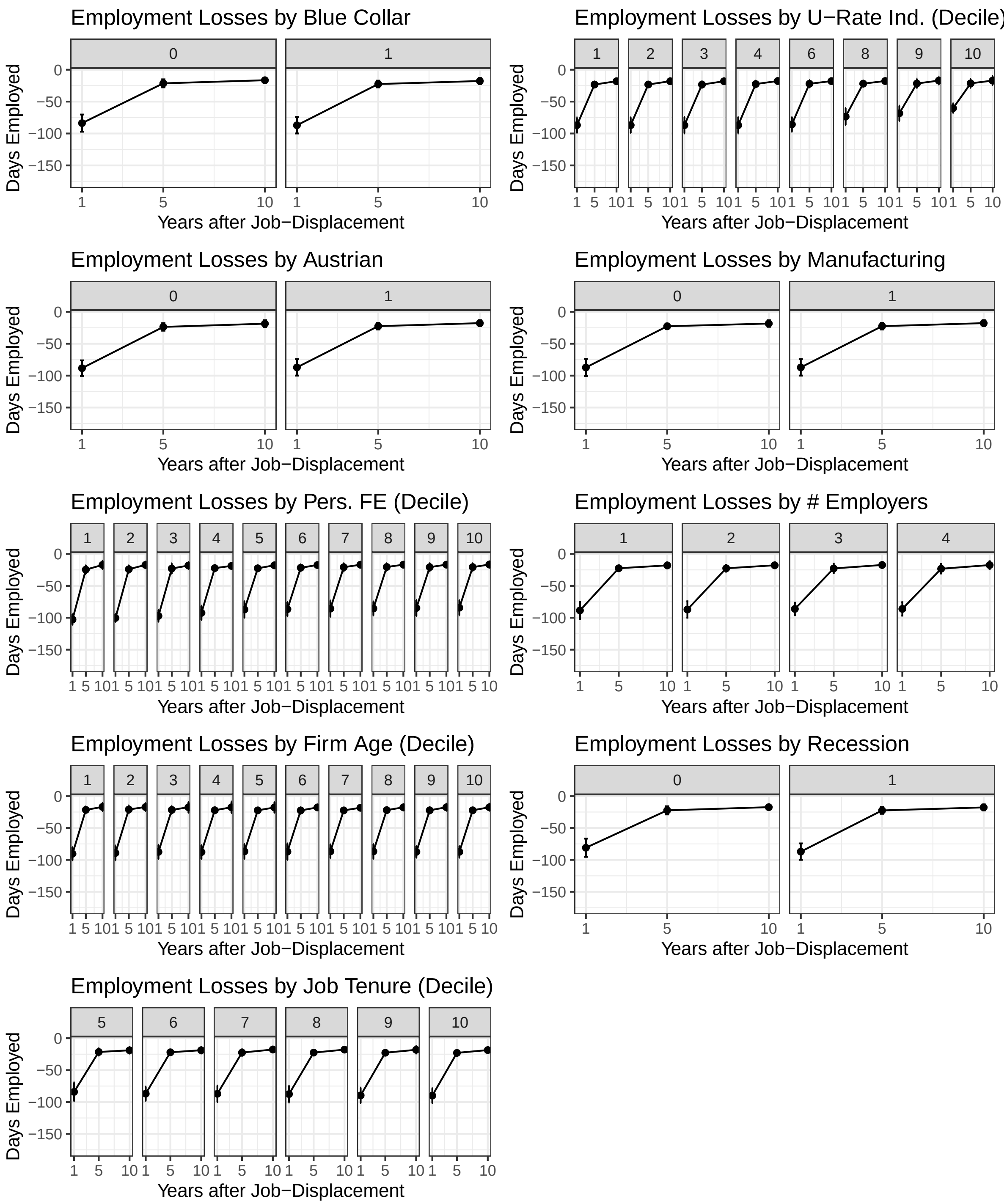


Figure 32: GRF estimates with 95% CI of employment losses by partitioning variables. All other variables are set to their median values. The boxplots present the distribution of the partitioning variable in the dataset (for the details see Subsection B of Section II).

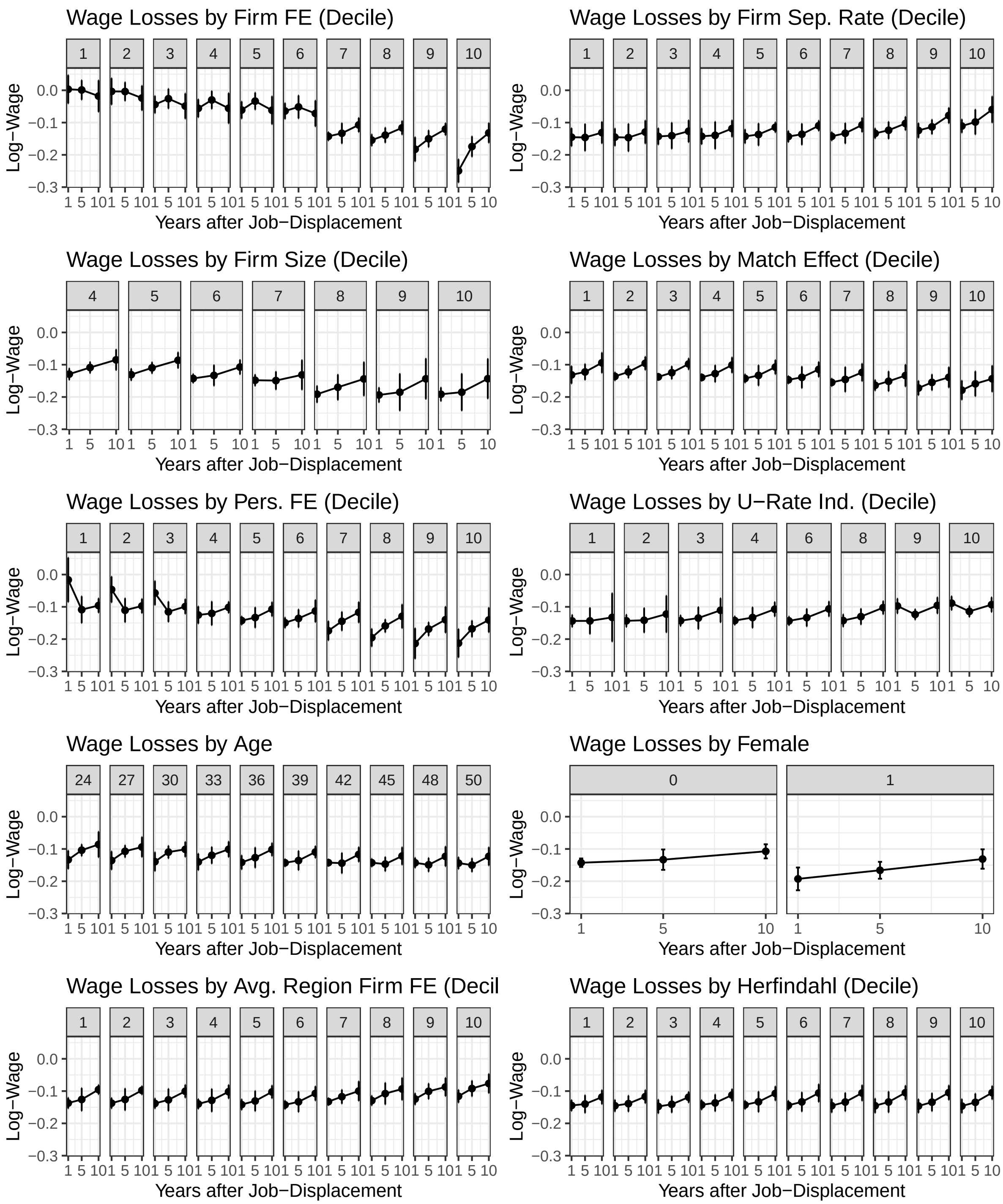


Figure 33: GRF estimates with 95% CI of wage losses by partitioning variables. All other variables are set to their median values. The boxplots present the distribution of the partitioning variable in the dataset (for the details see Subsection B of Section II).

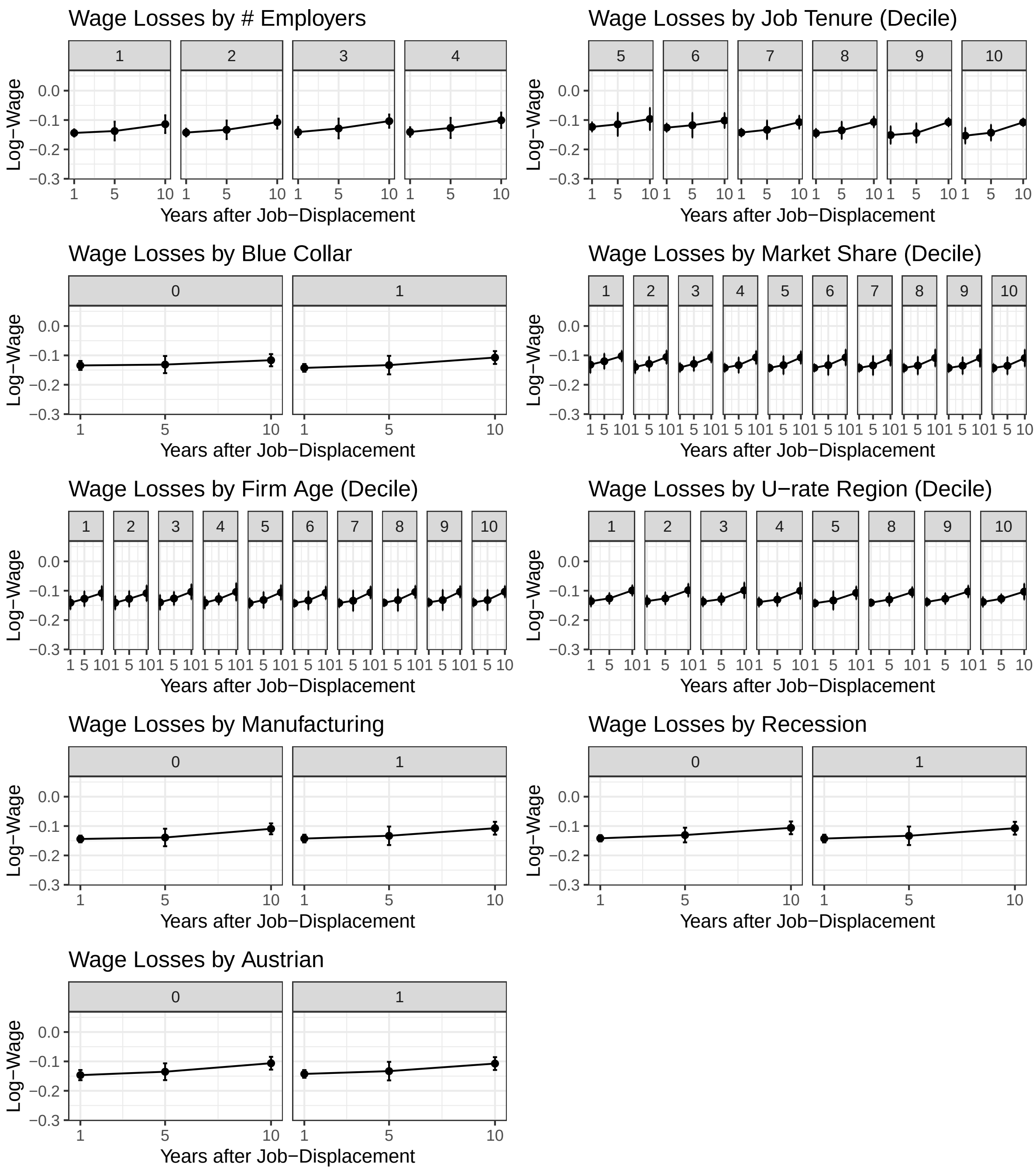


Figure 34: GRF estimates with 95% CI of wage losses by partitioning variables. All other variables are set to their median values. The boxplots present the distribution of the partitioning variable in the dataset (for the details see Subsection B of Section II).

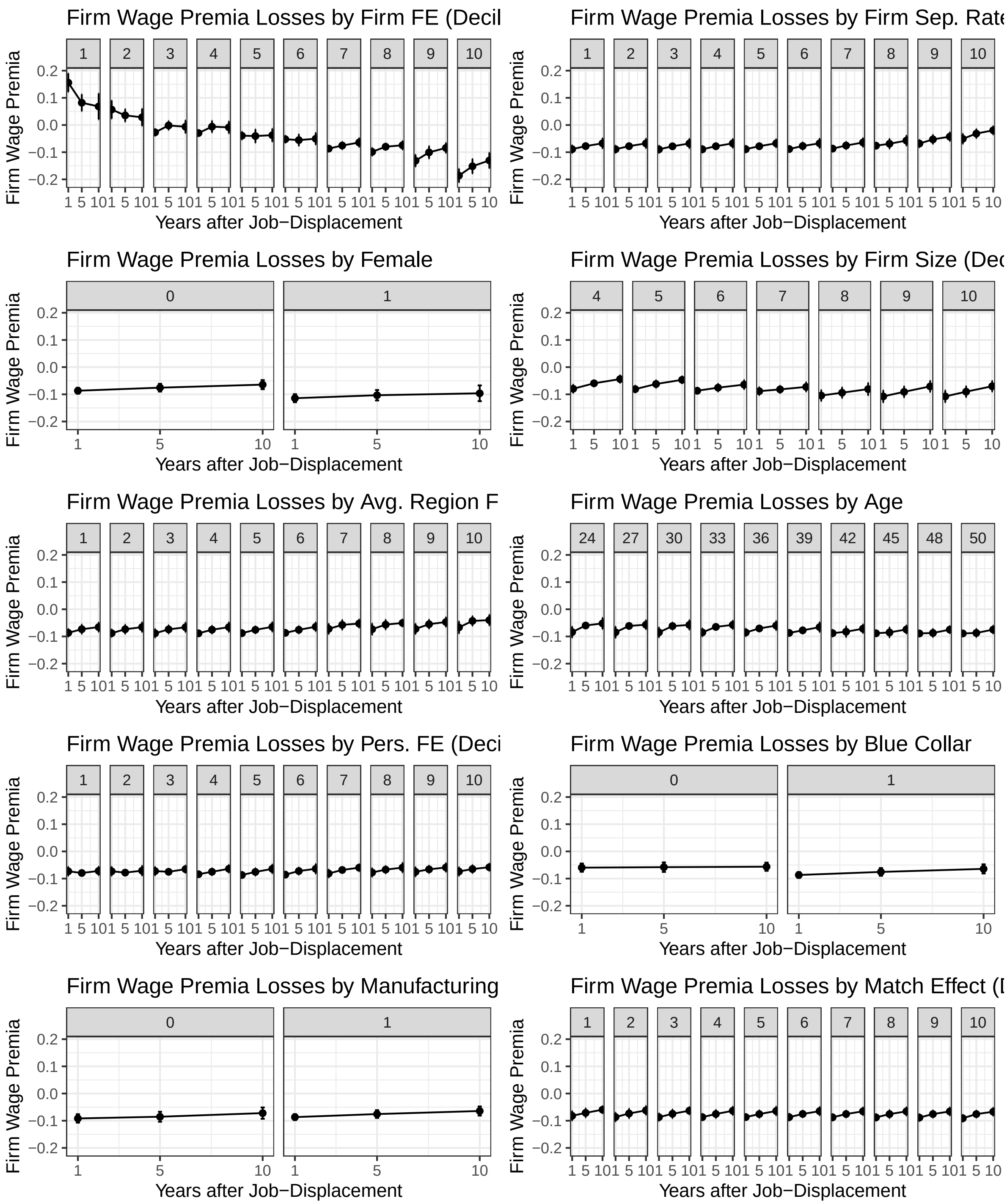


Figure 35: GRF estimates with 95% CI of firm wage premia losses by partitioning variables. All other variables are set to their median values. The boxplots present the distribution of the partitioning variable in the dataset (for the details see Subsection B of Section II).

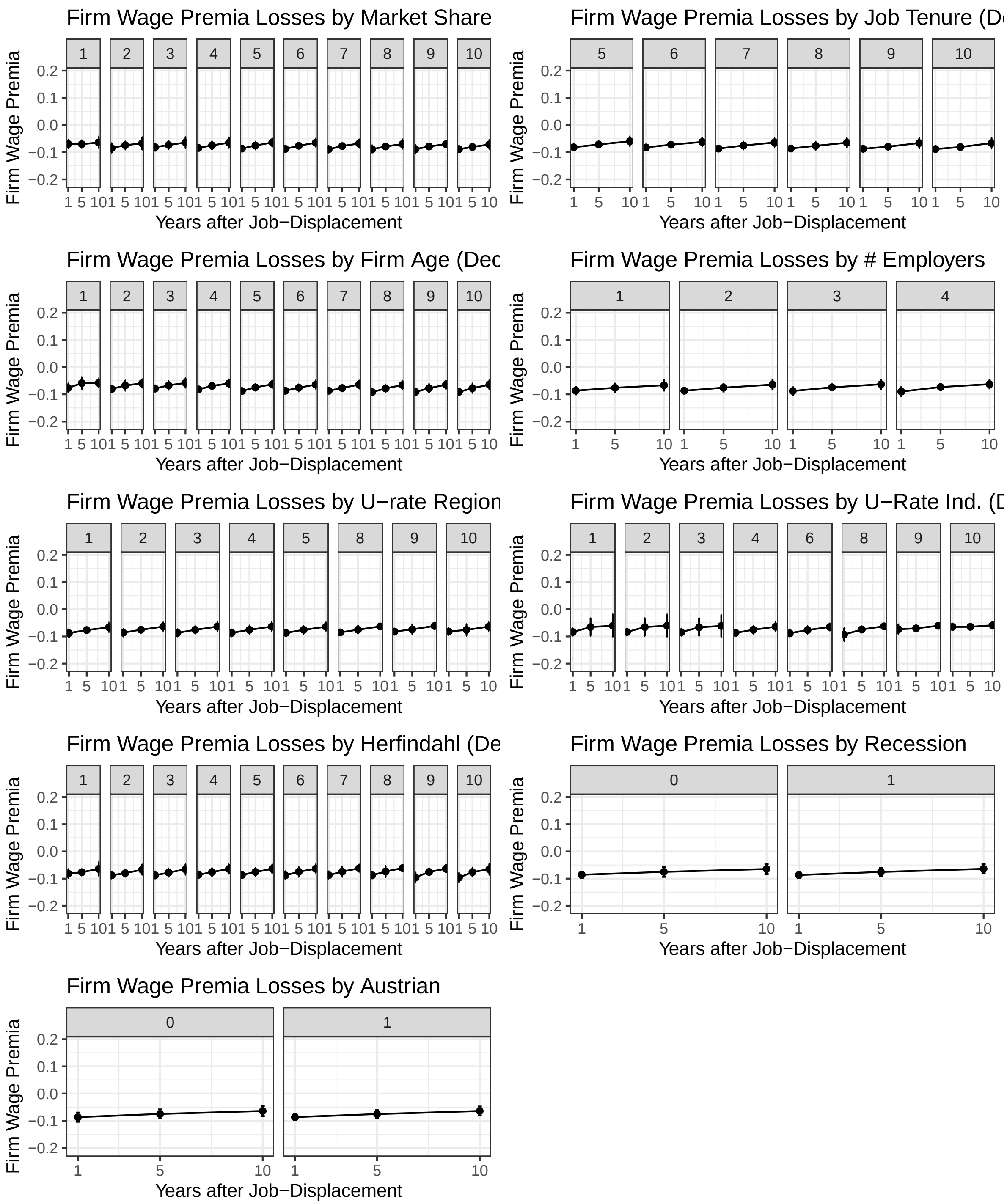


Figure 36: GRF estimates with 95% CI of losses in earnings by partitioning variables. All other variables are set to their median values. The boxplots present the distribution of the partitioning variable in the dataset (for the details see Subsection B of Section II).

K. Partial Dependence Plots

One potential criticism of the partial effects is that an individual, with median characteristics might not be representative for the whole population, and those effects might be very different from median realizations. To tackle this critique, we use partial dependence plots proposed by Friedman (2001) to better understand how a single variable affects *on average* the earnings losses in the sample. This approach consists in estimating the earnings losses for each individual by changing the value of one variable $z^k = \overline{z}$, while holding all other characteristics constant at their empirical values $\mathbf{z}^{-k}$. The counterfactual outcomes are then obtained by averaging over the sample distribution $F(\mathbf{z}^{-k})$. Formally we compute:

$$\mathbb{E}_{\mathbf{z}_{-k}} \hat{\tau}(z_k = \overline{z}; \mathbf{z}_{-k}) = \int \hat{\tau}(z^k = \overline{z}; \mathbf{z}^{-k}) \mathrm{d}F(\mathbf{z}^{-k}), \tag{19}$$

which in our application can be estimated on our training set: $\frac{1}{N}\sum_{i=1}^{N} \hat{\tau}(z^k = \overline{z}; \mathbf{z}_i^{-k})$. Figures 37 – 44 depict partial dependence plots of losses in earnings, employment, wages, and firm premia for different deciles of all partitioning variables.[35] All the main findings from the previous exercise preserve, with the difference, that the level of effects in the partial dependence plots is slightly shifted upwards. This is because earnings losses are higher for the average individual compared the to a worker with median characteristics.

[35] In this exercise due to a very high memory consumption, we needed to decrease the number of trees to 2,000 instead of initial 10,000.

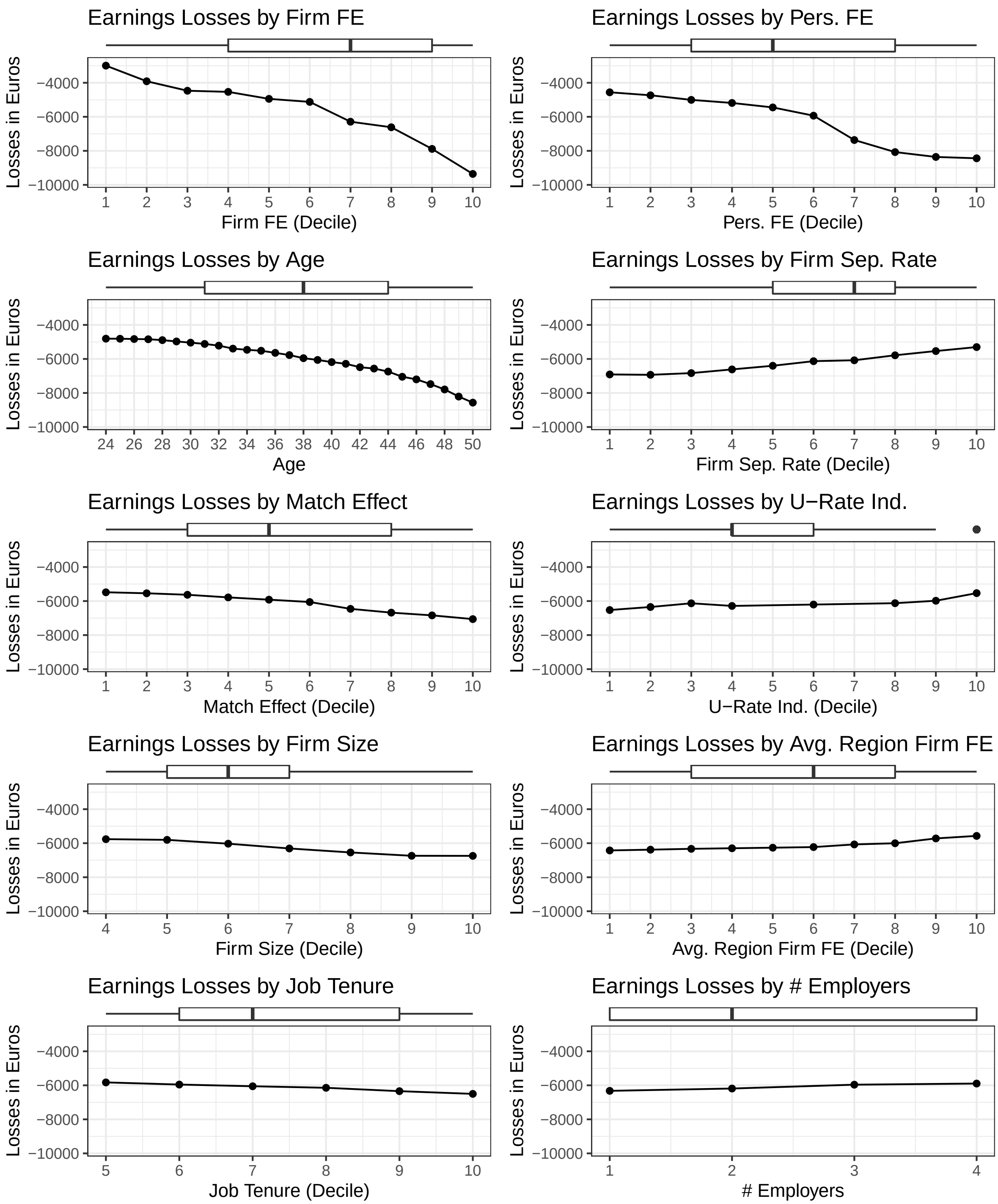


Figure 37: Partial dependence plots of earnings losses (part I). The boxplots above figures represent the distribution of the partitioning variable in the restricted dataset (for the details see Subsection B of Section II).

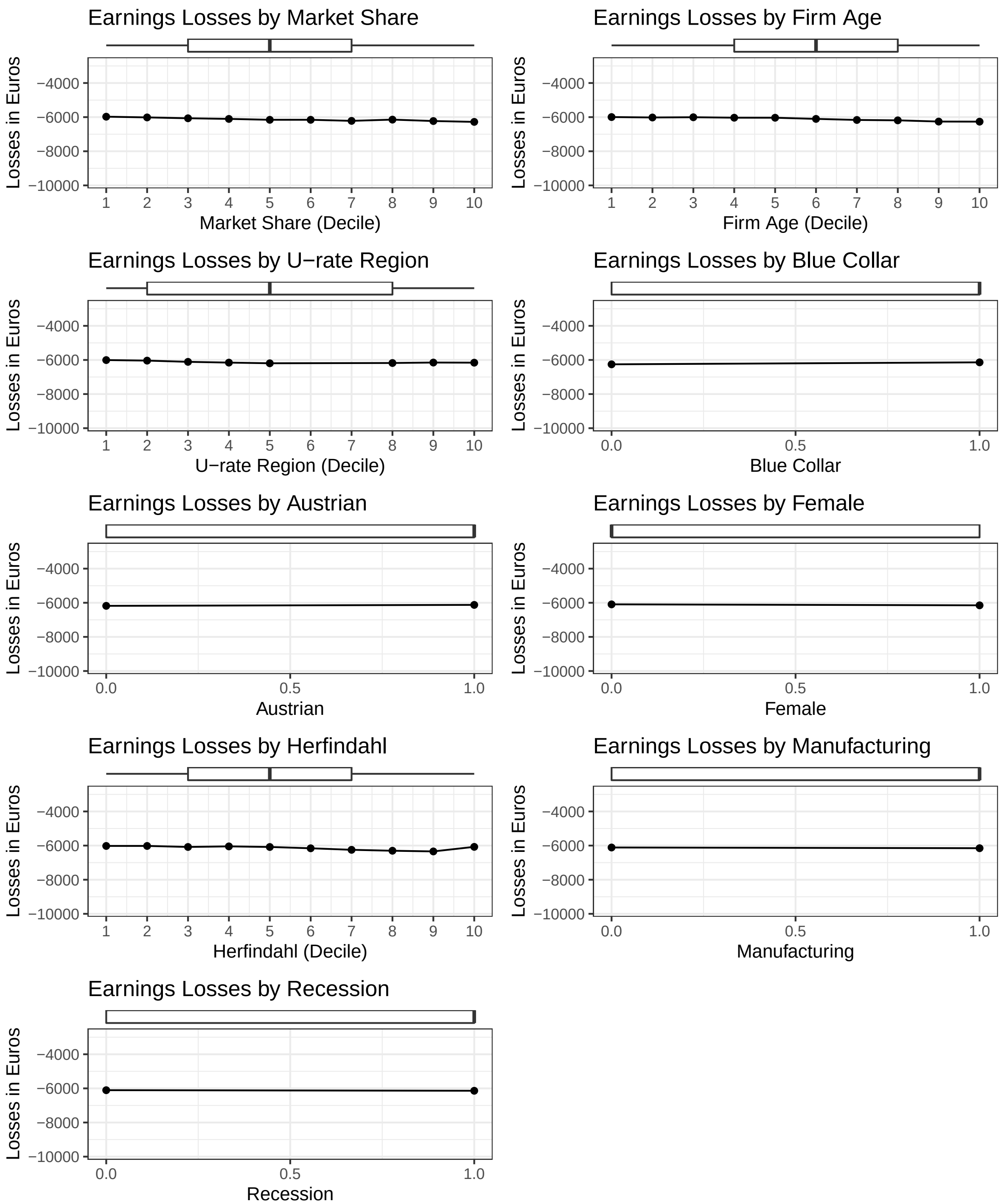


Figure 38: Partial dependence plots of earnings losses (part I). The boxplots above figures represent the distribution of the partitioning variable in the restricted dataset (for the details see Subsection B of Section II).

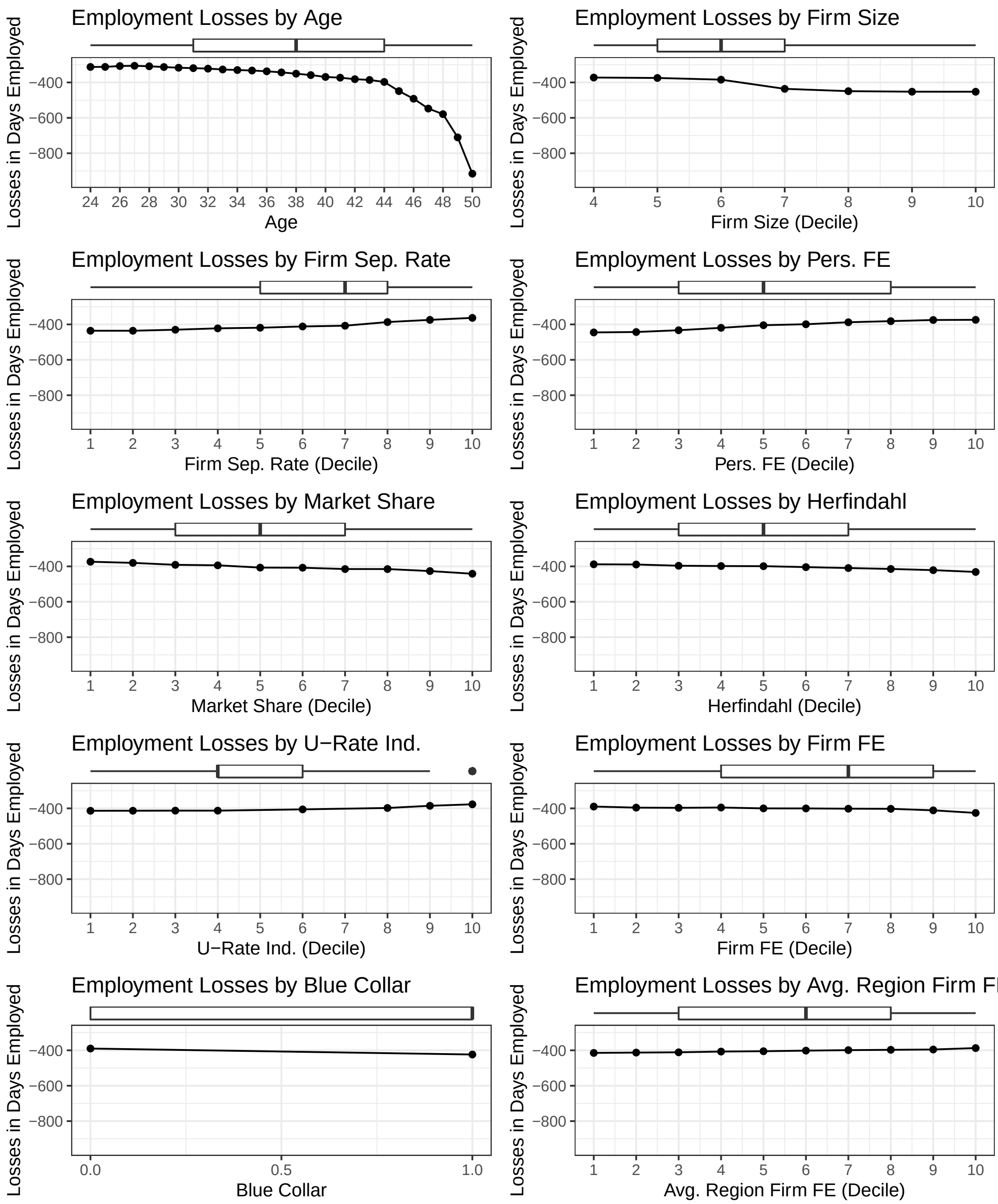


Figure 39: Partial dependence plots of employment losses (part I). The boxplots above figures represent the distribution of the partitioning variable in the restricted dataset (for the details see Subsection B of Section II).

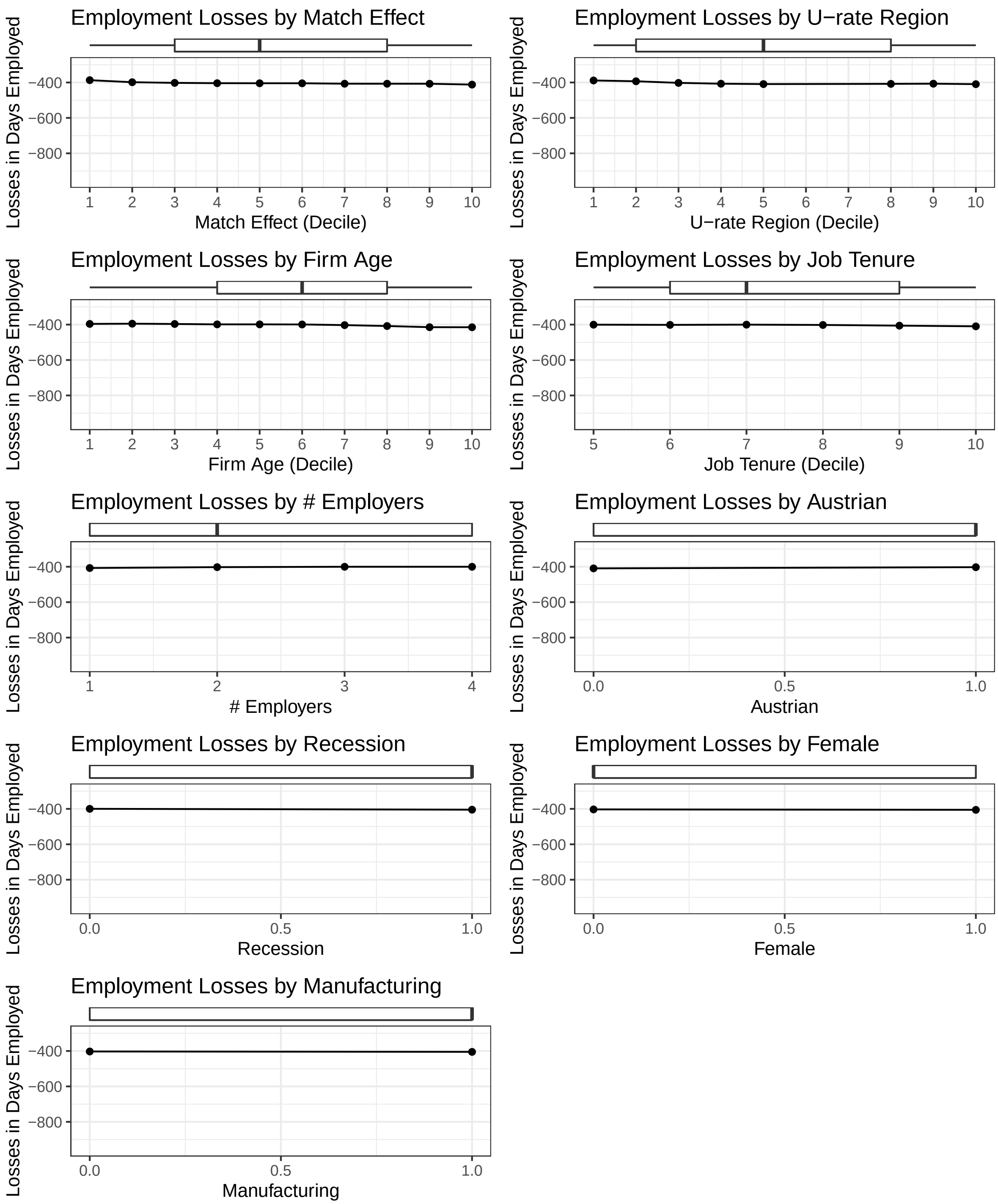


Figure 40: Partial dependence plots of employment losses (part II). The boxplots above figures represent the distribution of the partitioning variable in the restricted dataset (for the details see Subsection B of Section II).

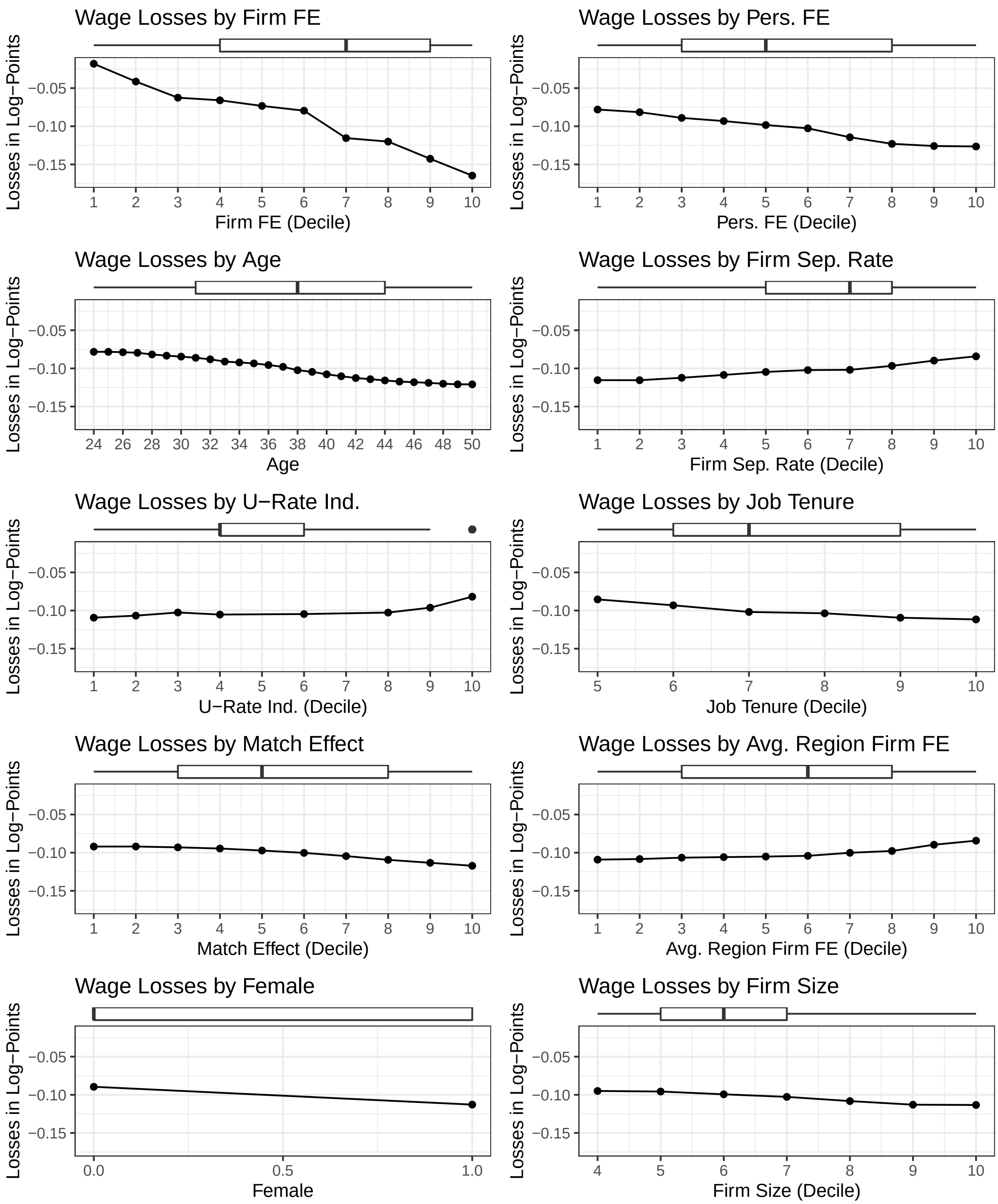


Figure 41: Partial dependence plots of wage losses (part I). The boxplots above figures represent the distribution of the partitioning variable in the restricted dataset (for the details see Subsection B of Section II).

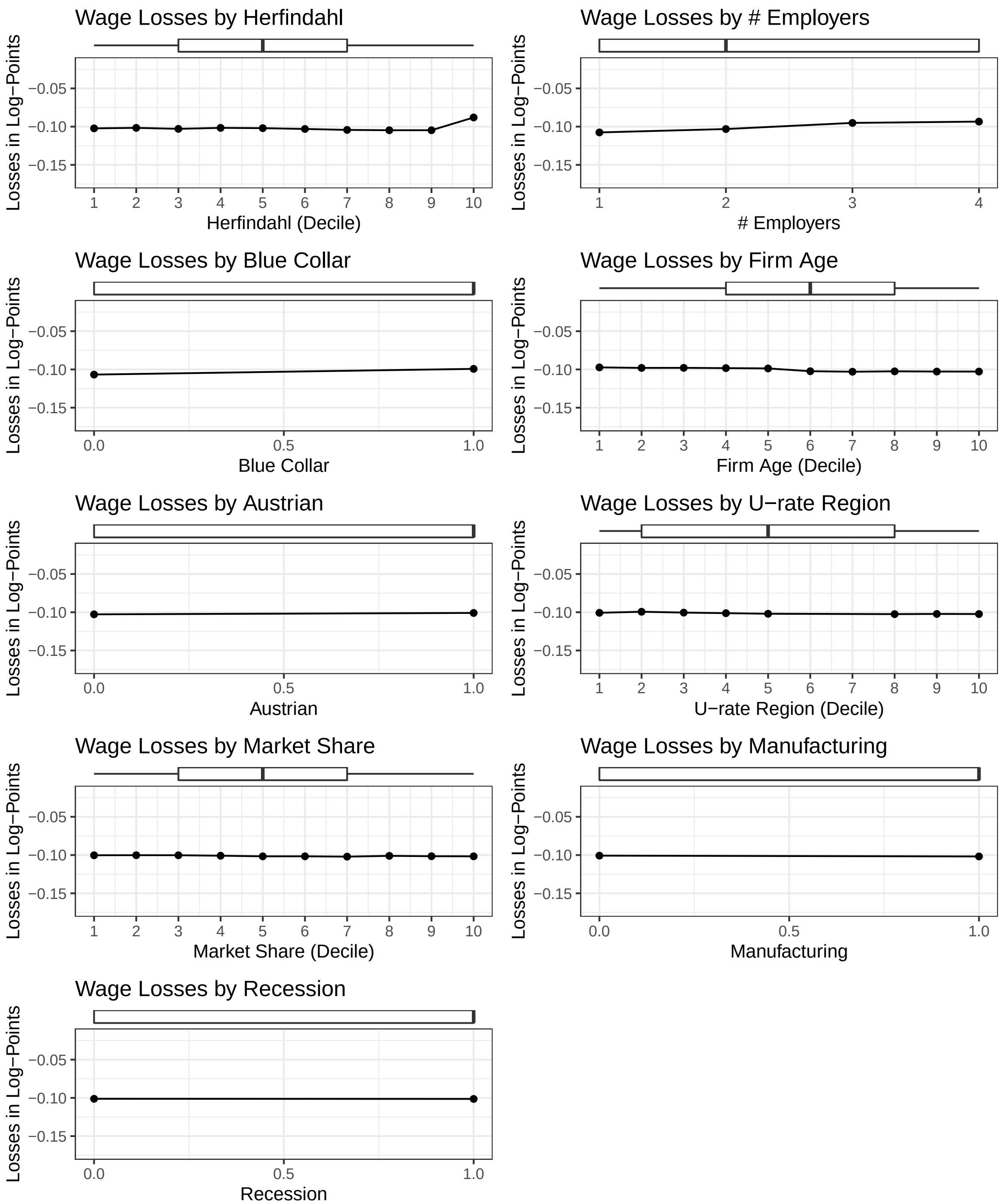


Figure 42: Partial dependence plots of wage losses (part II). The boxplots above figures represent the distribution of the partitioning variable in the restricted dataset (for the details see Subsection B of Section II).

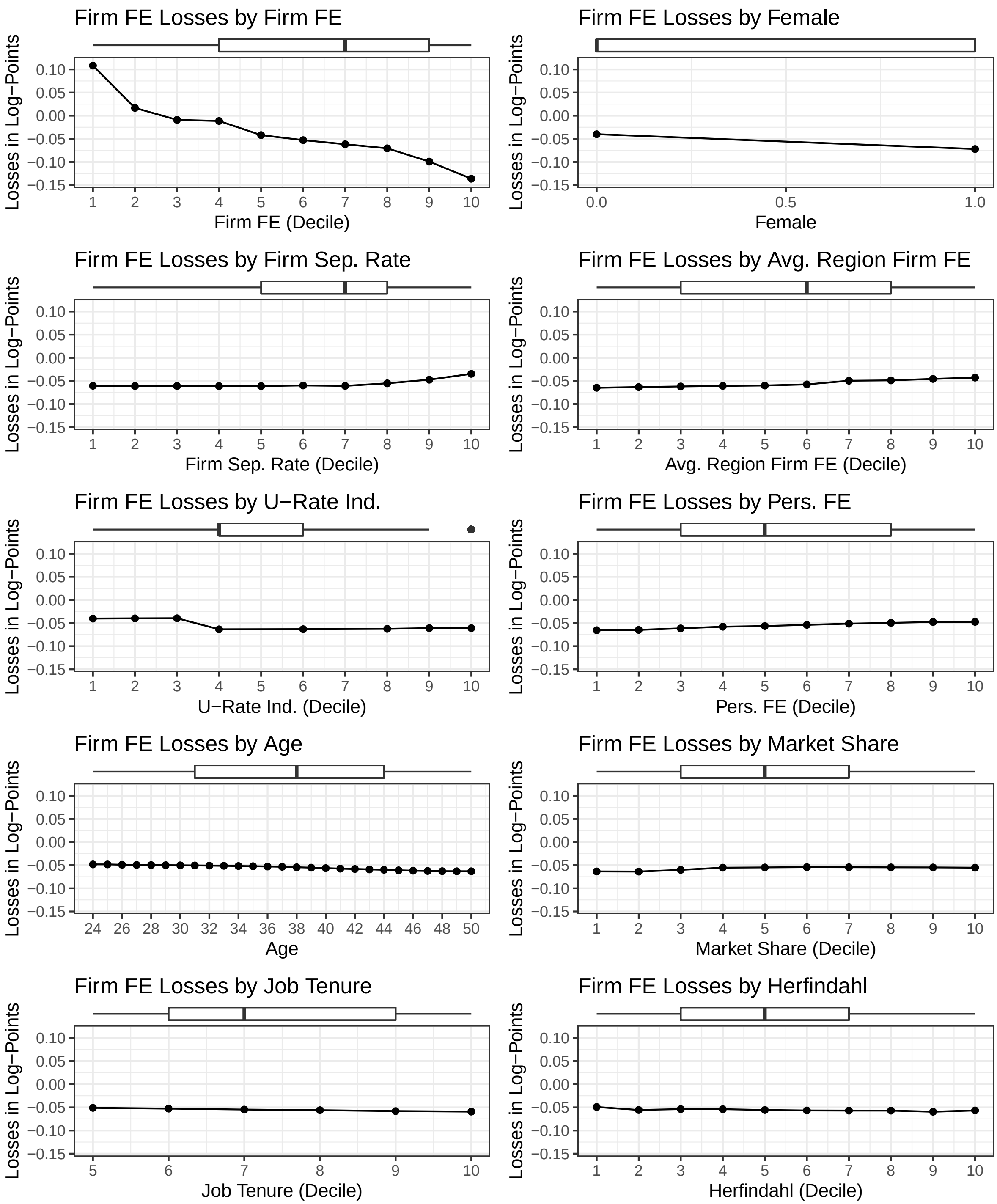


Figure 43: Partial dependence plots of the average 11-year firm wage premia losses (part I). The boxplots above figures represent the distribution of the partitioning variable in the restricted dataset (for the details see Subsection B of Section II).

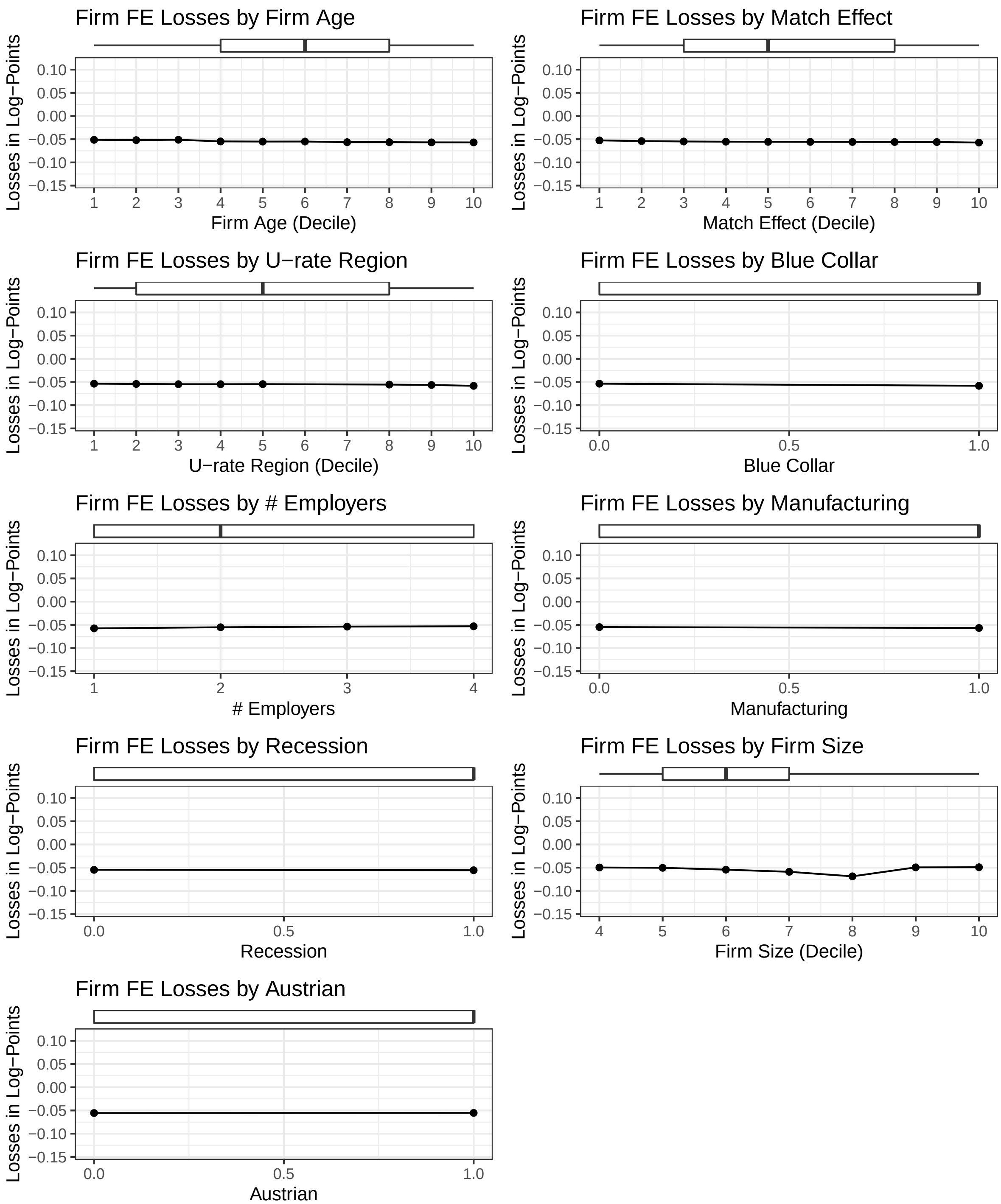


Figure 44: Partial dependence plots of firm wage premia losses (part II). The boxplots above figures represent the distribution of the partitioning variable in the restricted dataset (for the details see Subsection B of Section II).

L. Descriptive statistics

	Expansion	Recession
Blue Collar	0.60	0.59
Manufacturing	0.54	0.56
Austrian	0.72	0.75
Female	0.41	0.40
Age	37.80	37.68
Job Tenure	2,309.49	2,408.91
# Firms	2.51	2.39
Match Effect	-0.0005	0.0005
Firm FE	61.12	61.27
Pers FE	46.12	47.07
Firm Size	220.57	209.96
Firm Age	19.20	19.41
Market Share	0.001	0.0004
Firm Sep. Rate	0.33	0.33
Avg. Region Firm FE	50.38	50.71
Herfindahl	0.001	0.001
U-rate Region	0.10	0.10
U-Rate Ind.	0.11	0.11

Notes: Table shows the compositional differences between displaced workers in recessiosn and expansions. See text for details.

Table 11: Averages of partitioning variables for workers displaced during expansion and recession.

	Male	Female
Blue Collar	0.63	0.53
Manufacturing	0.61	0.46
Austrian	0.75	0.72
Recession	0.72	0.72
Age	37.78	37.62
Job Tenure	$2,398.10$	$2,356.16$
# Firms	2.49	2.33
Match Quality	-0.01	0.01
Firm FE	68.77	50.07
Pers FE	57.61	30.82
Firm Size	213.29	212.38
Firm Age	19.31	19.42
Market Share	0.001	0.0002
Firm Sep. Rate	0.33	0.34
Avg. Region Firm FE	50.62	50.62
Herfindahl	0.001	0.001
U-rate Region	0.10	0.10
U-Rate Ind.	0.12	0.10

Table 12: Averages of partitioning variables for female and male displaced workers.

M. Heterogeneous Treatment Effects: A Subgroup Analysis Approach

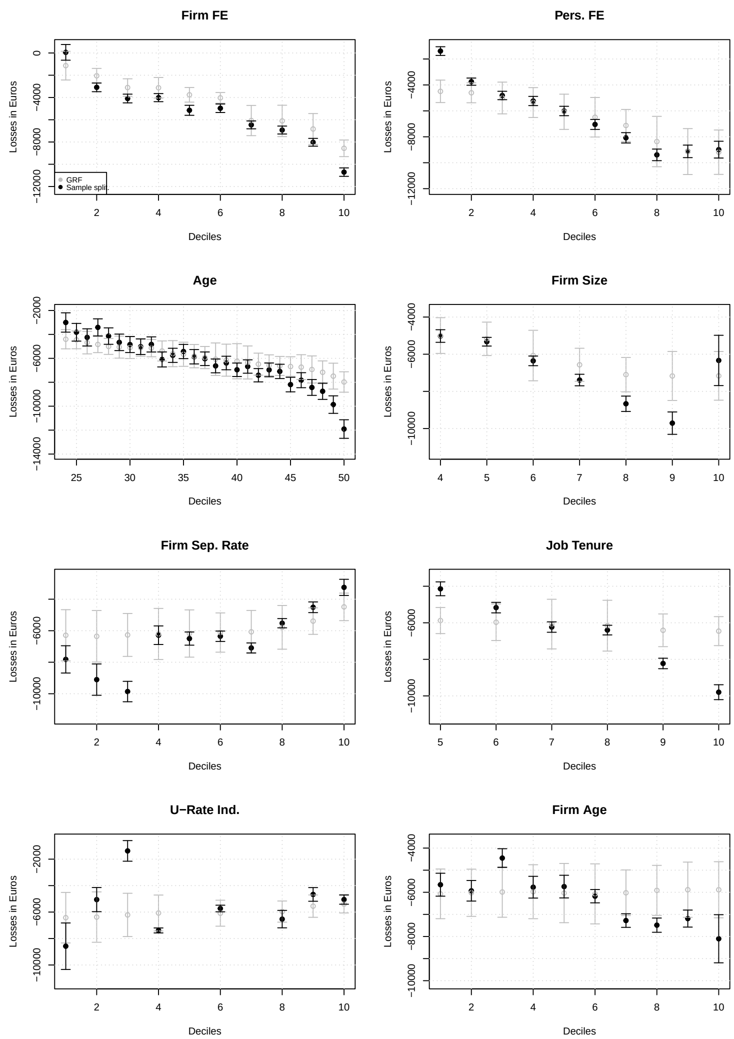


Figure 45: Earnings losses by different values of four selected variables (part I). Black and grey dots show earnings losses implied by sample splitting and the random forest, respectively.

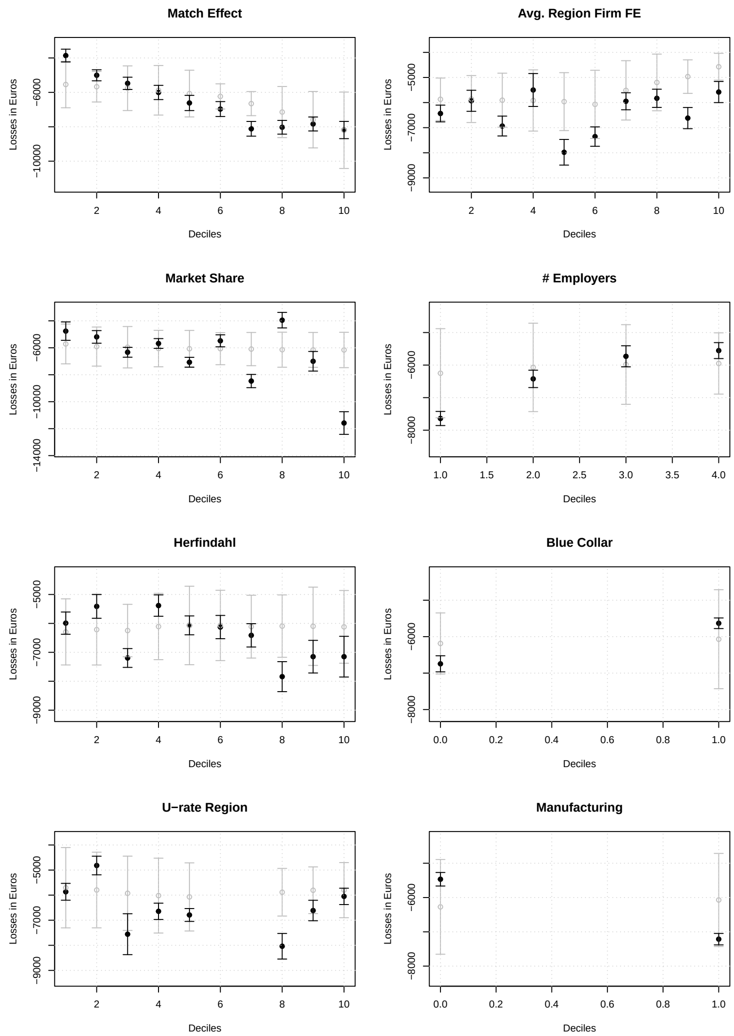


Figure 46: Earnings losses by different values of four selected variables (part II). Black and grey dots show earnings losses implied by sample splitting and the random forest, respectively.

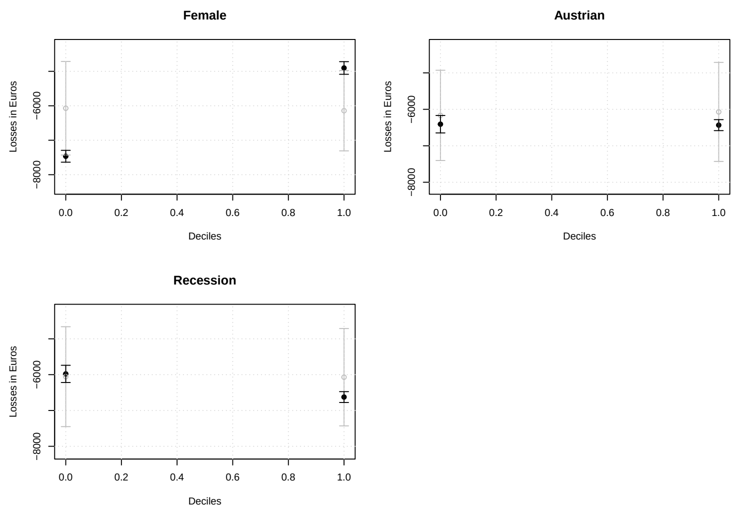


Figure 47: Earnings losses by different values of four selected variables (part III). Black and grey dots show earnings losses implied by sample splitting and the random forest, respectively.