Discussion Paper Series – CRC TR 224

Discussion Paper No. 131
Project A 03

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

Andreas Gulyas [1]
Krzysztof Pytka [2]

May 2020
(First version : October 2019)

[1] University of Mannheim, andreas.gulyas@uni-mannheim.de
[2] University of Mannheim, pytka@uni-mannheim.de

Funding by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) through CRC TR 224 is gratefully acknowledged.

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

Andreas Gulyas[†] Krzysztof Pytka[‡]

May 13, 2020

Abstract

We implement a generalized random forest (Athey *et al.*, 2019) to a difference-in-difference setting to identify substantial heterogeneity in earnings losses across displaced workers. Using administrative data from Austria over three decades we document that a quarter of workers face cumulative 11-year losses higher than 2 times their pre-displacement annual income, while almost 10% of individuals experience gains. Our methodology allows us to consider many competing theories of earnings losses. We find that the displacement firm's wage premia and the availability of well paying jobs in the local labor market are the two most important factors. This implies that earnings losses can be understood by mean reversion in firm wage premia and losses in match quality, rather than by a destruction of firm-specific human capital. We further show that 94% of the cyclicality of earnings losses is explained by compositional changes of displaced workers over the business cycle.

Keywords: Job displacement, Earnings losses, Causal machine learning

*Funding by the German Research Foundation (DFG) through CRC TR 224 (Project A3) is gratefully acknowledged. Johannes Gönsch and Michaela Paffenholz 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, and the Causal Machine Learning Workshop St. Gallen for helpful comments. Further results can be explored interactively using the companion web applet available at: `https://gulyas-pytka.app/earnloss`.

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

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

I. Introduction

A sizable literature documents large and long lasting consequences of job loss. Workers displaced during a mass layoff experience significant losses in annual income lasting over 15 to 20 years.[1] The sources behind the persistent earnings losses are still an open question in labor market research. Several channels have been put forward as explanations. For instance, two recent papers study the role of losses in employer-specific wage components and arrive at different conclusions. Schmieder *et al.* (2020) argue that they play an important role in Germany, whereas Lachowska *et al.* (2018) show that lost firm-specific human capital is more important for displaced workers in Washington during the financial crisis. Prior research has shown that long tenure in prior job, industry, and occupation lead to substantially larger losses, providing evidence that job-specific skills are lost with displacements.[2] Jarosch (2015) highlights the role of lost job security as another potential important channel. A particular challenge for identifying the contribution of these channels is that displaced workers differ among many correlated dimensions. Better paying or more stable firms might have longer tenured workers or specific workers sort into different types of firms.[3] Ideally, one needs to identify how certain variables affect earnings losses while holding confounding factors constant.

To resolve these challenges, we implement a generalized random forest (Athey *et al.*, 2019) to a difference-in-difference setup and estimate how the causal effect of job displacement varies with a large number of observables. The advantage compared to the prior literature in our view is that we do not rely on any modeling or functional form assumptions as in structural models, nor do we have to study losses within a particular reduced form model.[4] Instead, we draw upon the universe of male Austrian social security records over three decades, and use the variation in earnings losses by worker and job characteristics to identify the driving forces in a theory-agnostic and data-driven approach.[5] First, this enables us

[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] Jacobson *et al.* (1993), Neal (1995).

[3] *E.g.,* Card *et al.* (2013), Song *et al.* (2018).

[4] Jarosch (2015), Krolikowski (2017), Jung and Kuhn (2018) study the sources of earnings losses in structural search models, whereas Lachowska *et al.* (2018) and Schmieder *et al.* (2020) study earnings losses within the Abowd *et al.* (1999) framework.

[5] Earnings losses of job loss in the context of Austria have been studied before by e.g. Schwerdt *et al.* (2010) and Ichino *et al.* (2017). Halla *et al.* (2018) show limited increases of spousal labor supply after husbands' job displacement in Austria.

to document substantial heterogeneity in the individual cost of job displacement. To the best of our knowledge, this is the first paper documenting the whole distribution of the causal cost of job loss.[6] Second, we can jointly test which of the most prominent hypotheses about earnings losses are the most relevant. Contrary to traditional subgroup analysis, where subgroups differ in many characteristics, our algorithmic approach lets us study the channels holding confounding factors constant.

The machine-learning procedure estimates the causal cost of job loss non-parametrically as a function of observables. This is achieved by exploring possible data splits and choosing those ones for which between-group differences in earnings 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.[7] Another 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.

With the trained random forest, we can estimate the causal cost of job displacement for each individual in our sample. While median cumulative earnings losses in the 11 years following separation amount to €61,900, almost 1/5 of our sample incur losses of above €100,000. Close to 10% of individuals even gain in terms of earnings from displacement. High absolute losses translate to high percentage losses in terms of pre-displacement income as they are highly correlated. We follow the typical sample restrictions in the literature in order to get a clean identification of the displacement event, which involves conditioning on larger firms and long-tenured workers. A long-standing concern is whether these sample restrictions select individuals that are bound to experience large earnings losses. Is the estimated cost of job loss thus representative for the whole population? We use the random forest to predict earnings losses for the population that fails to meet the sample selection criteria and show that they are surprisingly similar.

We construct 15 variables that are associated with the most prominent theories discussed in the literature as plausible explanatory factors. We seek to understand how earnings losses

[6] Guvenen *et al.* (2017) study the heterogeneity in long-term earnings losses by age and income group from at least one year out of work in the US.

[7] 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.B in greater detail.

vary with these channels and to identify which of these are the most important ones. We consider theories that tie earnings losses to losses in job-specific human capital through an inclusion of job tenure. To quantify the effects of lost firm rents we identify firm wage premia through firm fixed effects in a Mincerian wage regression following Abowd *et al.* (1999). We also use this firm pay measure to construct a (leave-out) mean firm pay measure at the local labor market level. In addition we study whether losses originate through lost job security and include the average yearly firm-level separation rate in the five years prior to the mass layoff as an explanatory factor. Another hypothesis we consider relates earnings losses to losses of particularly good matches. Since mass layoffs at large firms can have regional spillover effects (Gathmann *et al.*, 2018), we also include firm size and the Herfindahl-Hirschmann index of labor-market concentration as explanatory variables. We further analyze the cyclicality of earnings losses by studying how the cost of job loss varies with regional and industry unemployment rates at the time of separation, as well as the state of the aggregate economy.

A contribution to the existing literature is that our methodology allows us to jointly compare all those theories. We show that from all these variables, firm wage premia is by far the most important. We arrive at this conclusion through three different analyses. 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. Not only is the pre-displacement firm wage premium by far the most important variable, but the second most important is the average firm-pay premium in the region. We further conduct a variance decomposition and show that the variation in firm wage premia alone explains 63% of the overall variation in earnings losses across individuals. Third, to understand the contribution of each variable individually, we compute how earnings losses change by varying one channel at a time, holding all other variables constant at their median. This reinforces the importance of lost firm premia in understanding earnings losses, which shows the steepest slope in losses. Earnings losses rise from €21,000 for workers employed at the lowest paying decile of firms to almost five times as much for the highest paying firms. We identify strong interactions effects with the availability of well-paying jobs in the region. Workers employed at low-paying firms, but located in regions with many well-paying jobs do not face any earnings losses. The changes in firm wage premia are driven by a mean-reversion pattern. Workers with above median firm wage premia lose in terms of firm pay, whereas the other half gain in terms of firm fixed effects. We also find steep slopes in losses for match quality and worker's age, but the two operate through different channels. While older workers face negligible wage changes but high employment losses, earnings losses for workers with high match quality 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 and match quality, 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 that earnings losses feature strong cyclicality.[8] 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 94% of the cyclicality. With our machine-learning approach we estimate the impact of a recession holding worker and job characteristics constant and we do not find that earnings losses vary significantly over the business cycle. This highlights the importance of the ability of our machine-learning approach to hold confounding factors constant.

Policy makers often respond to mass layoffs with special programs aiming to mitigate their consequences. We show how our estimates can be used to derive a simple tree-based decision rule to target individuals with above median losses using variables that can be easily observed and understood by everyone. The endogenously generated rule identifies individuals with high earnings losses with 75% accuracy and depends only on five criteria. The model suggest that targeting policies should be time-invariant and not focused on recessions.

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 (2017) who study the heterogeneous effects of youth employment programs, while Knaus *et al.* (2017) use a LASSO model to study treatment heterogeneity of job search programs.

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

[8] *E.g.* Davis and Von Wachter (2011), Schmieder *et al.* (2020).

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 2017. This data comprises day-to-day information on all jobs and unemployment spells covered by social security in Austria (Zweimüller *et al.*, 2009). It contains information on yearly earnings for each worker-establishment pair.[9] It further contains basic socio-demographic information at the worker level such as age, gender, occupation, and citizenship. Each firm (we use firm and establishment exchangeably from here on) has a unique identifier, which allows us to study changes in employer specific characteristics over time. At the establishment level we have data on the geographic location and a 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. A worker is considered displaced if he separated from his 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 it declined by more then 30 percent in size during year t. To avoid selecting volatile firms, we exclude firms that either grew rapidly the years before the mass layoff, or rebounded in size 3 years after the mass layoff event. More precisely, 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. In addition, to avoid mis-specifying mergers, outsourcing or firm restructures as mass layoffs, we compute a cross flow matrix for all firms in each year. We exclude all firms where more than 30 percent of its workforce ends up working for the same employer in $t+1$.[10] 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.

[9] We deflate all earnings to 2017 level using the CPI index provided by the Austrian Statistical Agency.
[10] For an in-depth discussion see Hethey-Maier and Schmieder (2013).

B. Construction of Sample

We focus on male workers only. In addition to enhancing comparability to previous studies that mainly study males, we do this because of two reasons. First, the higher labor force attachment of men reduces selection issues in our analysis. Second, the social security data only contains information on employment status, not on hours worked. Whereas during our study period, between 27-47 percent of women were working part-time, only between 4-11 percent of men were employed part-time.[11] We proceed by selecting everybody who is employed on the reference day of January 1st each year. This results in 45,689,628 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 8,276,796 person-year observations.

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. We identify 45,521 displaced (male) worker events between 1989 and 2007.

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 firm who did not separate.

Some workers disappear over time from our dataset. 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 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.

C. Propensity Score Matching

Non displaced workers may differ in many characteristics from the displaced workers. This is also confirmed by Table 1, which shows that displaced workers have more tenure, slightly higher earnings, are more likely to work in manufacturing and are employed in larger firms compared to non-displaced workers. In order to obtain a control group that is as similar

[11] Source: https://www.statistik.at/web_de/statistiken/menschen_und_gesellschaft/arbeitsmarkt/arbeitszeit/teilzeitarbeit_teilzeitquote/062882.html

	Displaced	Selected Control Group	Not Selected
Age	37.86	37.84	37.82
$\log(w_{t-1})$	4.65	4.66	4.71
$\log(w_{t-2})$	4.64	4.64	4.69
Job Tenure (in days)	2,560.97	2,537.33	2,661.38
Manufacturing	0.52	0.53	0.50
Firm Size	1,656.70	1,634.75	1,462.52
Obs	45,521	45,521	8,185,754

Table 1: Sample characteristics of workers

as possible to displaced workers, we use propensity score matching for the selection of our control group. In each year, for all workers satisfying the sample restrictions, we estimate the propensity to experience a displacement event as a function of observable worker and firm characteristics. Specifically, we use 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 as matching variables.[12] 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 16 in the appendix, which plots raw averages across these groups over time.

III. The Average Cost of Job Displacement

To estimate the cost of job displacement relative to the counterfactual of no displacement, we estimate the following regression model:

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

where D_i is an indicator equal to one for a displaced persons, t^* the displacement year and t the current year. We include year fixed effects γ_t to control for the evolution of the control group's earning. We also control for initial differences in earnings by including a

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

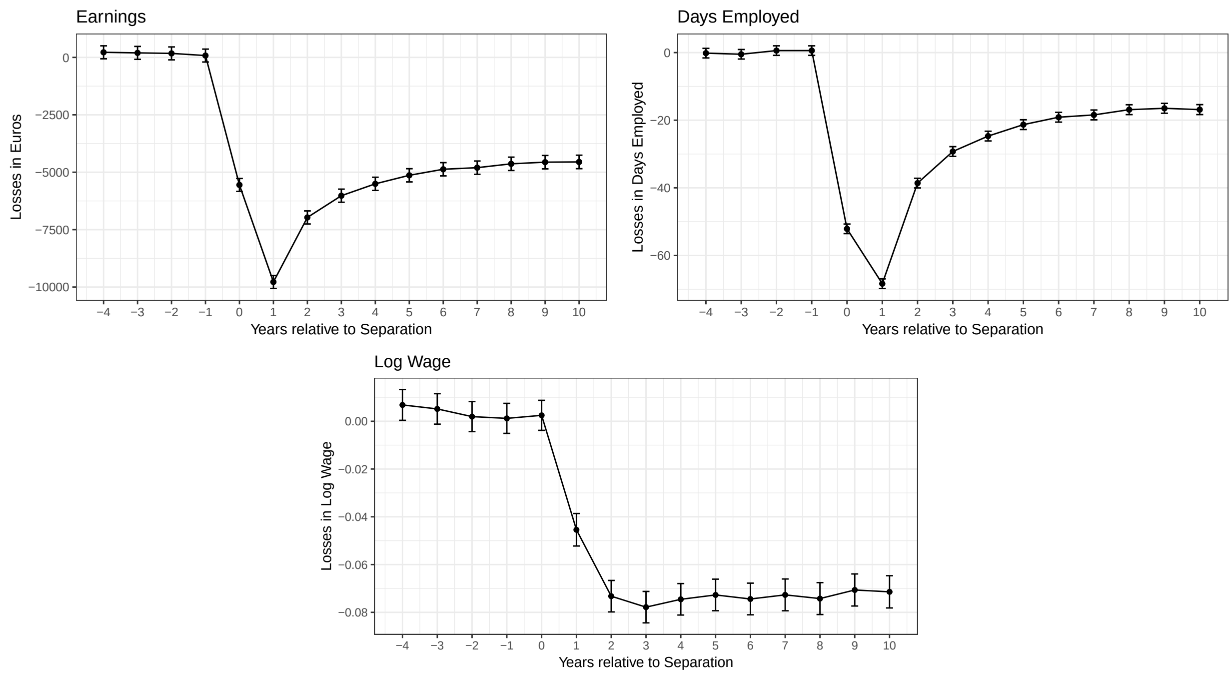


Figure 1: Earnings Losses of displaced workers - Eventstudy regression estimates of equation (1). Period 0 corresponds to the separation year. Earnings and days employed are computed for the whole year, log wages are computed as the log average daily wage from the employer on 1st January. Control group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

displacement dummy D_i.[13] Because of the propensity score matching, workers in the control group are very similar. δ_j consequently measures the change in workers earnings relative to the baseline year $t^* - 5$, after controlling for differences in initial earnings between the two groups.

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 1 shows the estimated causal effect of job-displacement for these three outcome variables. In the year after job displacement, earnings losses amount to approximately €10,000, 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 €5,000 yearly, or 13% in terms of pre-displacement earnings. Figure 1 further shows that this decline in earnings both stem from employment losses and declines in log-wages. In the two years after job displacement,

[13] Instead of controlling for differences in $t - 5$ with the group dummy D_i, we could have included worker fixed effects. The results are almost identical for the two cases, but because of the computational intensity of our machine learning algorithm we use the dummy variable approach already here.

employment losses amount to almost 50-70 days, but employment almost fully recovers after a couple of years. More strikingly, there is essentially no recovery in log wages. Displaced workers' wages decline by 7-8%, with no noticeable 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.

Throughout the paper we will mostly focus on the cumulative earnings losses in the 11 years following job displacement. These can be estimated either by summing appropriately over the Event Study coefficients ($\sum_{j=0}^{j=10} \delta_j$) from equation (1), or more compactly using the following difference-in-difference setup:

$$y_{it} = \tau \mathbb{1}(t \geq t^*) \times D_i + \theta D_i + \gamma_t + \epsilon_{it}, \tag{2}$$

where τ measures the average yearly cost of job loss, or if scaled by 11, the cumulative losses over 11 years. Table 2 shows the estimates from this regression for different specifications. Column (1) reports the estimates for equation (2) without any controls, in column (2) and (3) a polynomial in age and worker fixed effects are added. In all specifications, the yearly earnings losses amount to close to €5,900 per year, or close to €65,000 over 11 years. Because of the computational burden of the machine learning algorithm, we are 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.

IV. Identifying Sources – 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, very recently developed by Athey *et al.* (2019). There are several important advantages of our approach in comparison to traditional ones, *i.e.* estimating the homogeneous scarring effect or sample splitting to compute heterogeneous effects.

The average cumulative earnings losses after job displacement can be estimated directly from formula (2) using standard statistical methods. Nonetheless, this approach has some serious limitations. First and foremost, this type of the estimate provides the average treatment effect and does not give a gist on its underlying heterogeneity. If the earnings losses

	Dependent variable:		
	Yearly Income		
	(1)	(2)	(3)
$\hat{\tau}$	−5,850.6	−5,981.3	−5,952.0
	(55.2)	(33.0)	(32.1)
Worker FE		✓	✓
f(age)			✓
Observations	1,365,468	1,365,468	1,365,468
R^2	0.04	0.7	0.7
Adjusted R^2	0.04	0.7	0.7

Table 2: DiD Regression. Estimation results of Equation (2) for different sets of controls

are not uniform across individuals, then the individual scarring effects may be a far cry from the estimated average. This concern is of particular importance for policy makers, because accurate estimates of individual earnings losses are a prerequisite for designing optimal labor-market policies. In addition to this, sample restrictions typically applied in the literature, such as the focus on mass layoffs, age of workers, firm size, and job tenure, are the objects of discussion how the results can be extrapolated outside the restricted sample. Differences in the composition between the restricted sample and the general population may mechanically result in a situation in which the average effect identified using the training sample is not representable for the general population of the interest.

Heterogeneous treatment effects are typically identified through sample splitting and estimating the model on separate data bins. An application of this method can be illustrated by a study on heterogeneous scarring effects of one-year non-employment due to Guvenen *et al.* (2017). While this approach is absolutely correct for identifying the existence of heterogeneity, however, it does not tell us what are its driving channels. The reason for this is that the composition of individual characteristics may vary across data subsets in a similar fashion to differences between the restricted dataset and the general population, mentioned before. Consequently, it is impossible to disentangle the impact caused by the variable according to which the split has been made and the changes in the composition of the other variables.

Our approach consists in estimating the version of equation (2) 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 variables (henceforth called *partitioning variables*) for individual i and $\tau(\mathbf{z}_i)$ is a treatment effect for individual i. The functional specification of $\tau(\mathbf{z}_i)$ is assumed to be unknown. In theory, one can imagine computing (3) for all possible data bins, for each value of every partitioning variables interacted with other values of all other partitioning variables. That said, it would be very problematic for two reasons. First, it would mean that in our application there would be 1.04×10^{12} subsets.[14] This is much more than the number of observations, which makes the whole procedure computationally infeasible. Second, even if the computation of $\tau(\mathbf{z}_i)$ for each data subset were theoretically possible, it would not be recommended as the size of data subsets would be very small implying potentially low accuracy of the estimates. Therefore, we decide to use an algorithmic method to learn the pattern of $\tau(\mathbf{z}_i)$. This conditional average treatment effect is estimated by implementing a generalized random forest (Athey *et al.*, 2019), to a difference-in-difference setting. For every realization of $\mathbf{z}_i$ we derive weights which measure similarity of observations in terms of their cost of job loss. These weights then in turn are used in estimating $\tau(\mathbf{z}_i)$ with weighted least squares of equation (2). The weights are derived using a generalized random forest. The forest consists of many tree based models in which each tree is grown on a random subset of observations and partitioning variables. The algorithm explores possible data splits and chooses those ones for which between-group differences in the cost of job losses are the highest. By recursively splitting the dataset into smaller and smaller subsamples, the procedure endogenously partitions the observations into groups of workers with similar earnings losses. The weights measure how often individuals fall into the same partitions across all trees.

We proceed by describing the partitioning variables, which are associated with the most prominent channels from the earnings loss literature. Details on how we implement generalized random forests for our application can be found in subsection IV.B.

[14] Even if we would categorize every variable considered into two categories, we would still end up with 32,768 subgroups. Thus, even though we use the universe of social security records, we would have less than two displaced workers per group.

A. Partitioning Variables

The literature on earnings losses discusses a number of channels that could explain earnings losses.[15] We deliberately construct 15 partitioning variables that capture most of the channels that are widely considered to be important in explaining earnings losses.

One of the most prominent theories about earnings losses is based on the notion of firm-specific human capital. The idea is that workers accumulate job specific human capital over the course of their tenure at an employer. Earnings losses upon job loss therefore reflect the lost job-specific human capital that was embodied in the job. This channel is captured by the inclusion of job tenure as a partitioning variable. We in addition include the number of distinct previous employers (censored in 1984). Workers with higher past job mobility might be less prone to high earnings losses.

A growing literature shows that differences in firm pay is a significant driver of wage inequality (Abowd *et al.*, 1999; Card *et al.*, 2013, , among many others). Workers might fall off the job ladder and find reemployment at firms paying lower wages (Jarosch, 2015). Two recent working papers (Lachowska *et al.*, 2018; Schmieder *et al.*, 2020) highlight lost employer rents as an important reason for reduced earnings after job loss To understand how much of earnings losses might be attributed to losses in rents paid by firms, we include a firm fixed effect in a wage-regression following Abowd *et al.* (1999) (henceforth, AKM). Despite well known limitations about the structural interpretation of firm wage premia, the AKM model has become the workhorse model for empirically estimating the firm pay component.[16] 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[17] 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. We use the universe of male private sector employment spells in Austria from 1984 through 2017 in the estimation. In order to avoid endogeneity issues, we disregard individuals in our earnings loss sample from the computation of firm fixed effects.[18] This

[15] See Carrington and Fallick (2017) for a review.

[16] Eeckhout and Kircher (2011), Bagger and Lentz (2018), Hagedorn *et al.* (2017), Gulyas (2018).

[17] The dominant employer is selected based on the total earnings in calendar year t.

[18] 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.

implies that from 56,254,607 person-year observation, 2,427,901 are dropped. Table 10 in Appendix A shows that the variation in the firm fixed effect alone explains around 30% of the variation of the log wages. We cannot rule out that the mass layoff event nevertheless affected the estimation of the firm fixed effects through remaining workers who were not selected by the propensity score matching. Therefore, we will also consider a robustness exercise, in which we compute the AKM firm fixed effects only on observations before 1998, and we will restrict our earnings loss sample to years after 1998.

Davis and Von Wachter (2011) and Schmieder *et al.* (2020) document that earnings losses vary over the business cycle in the US and Germany. We capture the potential impact of the business cycle at the aggregate level by creating a dummy accounting for recession years according to the OECD definition. In addition, we consider the unemployment rate in the local labor market and industry of workers on our reference day. We use the NUTS-3 district (35 categories) and NACE level-1 industry classification (21 categories) of employers. The NUTS-3 classification follows European standards and is a sensible definition of a local labor market and is roughly comparable to US counties.

We are further interested how the availability of well-paying jobs affects earnings losses. We build on our firm wage premium estimates from the AKM regressions and compute the average firm premium in the region. 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.

Another hypothesis is that displaced workers lose a particularly good match. Some worker skills are only valued at particular firms and earnings losses might reflect that those skills are not demanded by other employers. To study this channel, we estimate the match effect of worker i employed at firm $J(i,t)$ as the residual term ϵ_{it} from the following regression:

$$\ln(w_{it}) = \alpha_i + \hat{\psi}_{J(i,t)} + \theta_t + f(age_{it}) + f(tenure_{it}) + \epsilon_{it}, \tag{5}$$

where $f(age_{it})$ and $f(tenure_{it})$ are cubic polynomials and $\hat{\psi}_{J(i,t)}$ is the estimated firm fixed effect from regression (4).

Recent evidence shows that many local labor markets are highly concentrated (Azar *et al.*, 2017). Since mass layoffs are conditional on large firms, the decline of these firms might have strong spill-over effects on the whole region (Gathmann *et al.*, 2018), further depressing outcomes for displaced workers. This effect is presumably stronger in large firms and concentrated labor markets. We therefore include firm size and construct the Herfindahl-

Hirschman index of labor market concentration, which is the standard concentration measure used in anti-trust cases. For each region and industry combination in each year, we compute $H_{yri} = \sum_j s_j^2$, where s_j is the employment share of firm j in year y, region r, and industry i.

Furthermore, we include a number of socio-demographic factors such as worker age, a dummy for Austrian citizenship, blue collar occupation, and manufacturing sector. These per-se might not be informative about potential channels, but we want to make sure that any of the identified effects are not originating simply through compositional differences in socio-demographic factors. This ability to control for observable differences is in fact one of the advantages of our machine-learning approach compared to simple subgroup analysis.[19]

To enhance interpretability, we categorize all continuous variables into deciles according to the overall distribution of Austrian male 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. Figure 17 in the appendix shows 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.

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

In the implementation part, we implement a generalized random forest Athey *et al.* (2019) to a difference-in-difference setup. Consequently, in our design we need to include time fixed effects to control for preexisting trends in worker earnings.[20] For this reason, we are not able to use the `grf` library off the shelf, because it supports only simple regression models with one regressor (least squares, IV, and quantile). The exact implementation is provided below.

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:

$$\mathbb{E}(\mathbf{x}_{it}' \varepsilon_{it} | \mathbf{z}) = \mathbf{0}_{18}, \tag{6}$$

[19] Schwerdt *et al.* (2010) for example shows that white collar workers face higher earnings losses in the short run.

[20] One might think that the problem could be addressed simply by using time fixed effects as additional partitioning variables. Nonetheless, this may give rise to detection of some spurious treatment effects. To illustrate this risk, think about estimating a model $y_{it} = \alpha + \gamma(\mathbf{z}_{it}) + \varepsilon_{it}$ for a data set drawn from a process with zero treatment effect, $\gamma = 0$ and a positive linear time trend for y_{it}. From this procedure we obtain a positive treatment effect for $t > t^*$.

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 (2), 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 the it-th observation to estimating the cost of job loss at $\mathbf{z}$, and estimating equation

$$\begin{aligned}(\tau(\mathbf{z}), \theta(\mathbf{z}), \boldsymbol{\gamma}(\mathbf{z})) \quad = \quad & \text{argmin}_{\{\tau,\theta,\gamma\}} \left(\frac{1}{NT}\sum_{i}^{N}\sum_{t}^{T}\alpha_{it}(\mathbf{z})\mathbf{x}'_{it}u_{it}\right)\left(\frac{1}{NT}\sum_{i}^{N}\sum_{t}^{T}\alpha_{it}(\mathbf{z})\mathbf{x}'_{it}u_{it}\right)', \\ & \text{s.t.}\forall_{i,t}u_{it} = y_{it} - \tau\mathbb{1}(t\geq t^*)\times D_i - \theta D_i - \gamma_t. \end{aligned} \tag{7}$$

Notice that the problem (7) 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.

B.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 (2) separately.[21]

The main difference of our procedure from the textbook one, which can be found in 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.[22] In our setup, we are interested in growing a tree that explores the underlying heterogeneity of earnings losses

[21]It is noteworthy that while all parameters from (2) are estimated, only the parameter of our interest, the scarring effect, is used in the splitting criterion (8).

[22]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 throuhg 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.

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 (2) on $\mathcal{P}_l$ and $\mathcal{P}_r$.

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

$$(\tau_l - \tau_r)^2 \frac{n_l \cdot n_r}{N^2}, \tag{8}$$

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 (8) 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, $\mathcal{P}$.

---

between partitions of individuals with different characteristics. For this reason, we adapt the criterion (8) 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 it 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: (1) chooses which variables are important and contribute to accounting for the heterogeneous scarring effects and which do not; (2) detects non-linear relationships between τ and $\mathbf{z}_i$; (3) detects interactions (including interactions of higher orders) between partitioning variables.

Figure 2 depicts a tree grown using the described algorithm. Every node is labelled with the average cumulative 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 in 11 years is equal to around €65,000. In the first iteration, the split that maximizes heterogeneity between groups goes according to the firm fixed effect. Individuals displaced from firms paying above

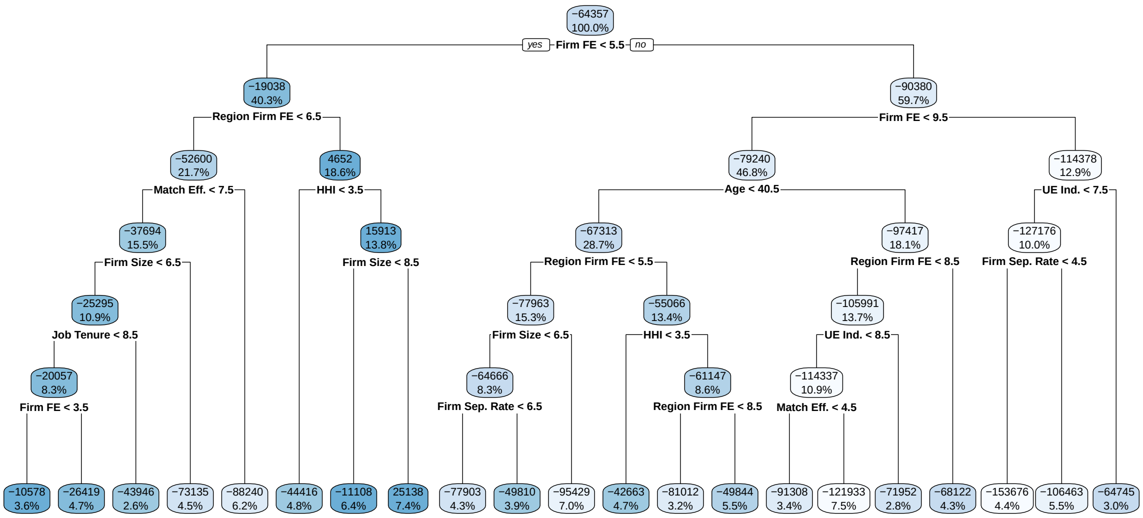


Figure 2: Heterogeneous treatment effect of Job Displacement. Tree was built with a CART algorithm using the relabeling strategy as described in the text. Min node size of 35000.

the median increase to €90,380, while the workers fired from firms paying below that level suffer losses lower than 20,000. Overall, the partitioning procedure generates 21 final nodes endogenously. The number of binary splitting conditions used to generate each leaf varies from 3 to 7. Already at this point, we can observe that high-income 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 work for firms with the highest wage premia and the separation rate below the median in industries with unemployment rate below 8th decile as it can be seen in the third leaf from the right.

B.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. A 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.[23]

B.3. 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:

$$\alpha_{it}^b(\mathbf{z}) := \begin{cases} \frac{1}{|L_b(\mathbf{z})|}, & \mathbf{z}_{it} \in L_b(\mathbf{z}) \\ 0, & \text{otherwise,} \end{cases} \tag{9}$$

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 (7) 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 (8). 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:

$$\frac{1}{N_{\mathcal{P}}} \sum_{(i,t)\in\mathcal{P}} \mathbf{x}_{it}' u_{it} = \mathbf{0}_{18}. \tag{10}$$

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

$$\tau_{C_k} \approx \tau_{\mathcal{P}} + \xi' \left(\frac{1}{N_{\mathcal{P}}} \sum_{(j,s)\in\mathcal{P}} \mathbf{x}_{js}\mathbf{x}_{js}' \right)^{-1} \cdot \frac{1}{N_{C_k}} \sum_{(i,t)\in C_k} \mathbf{x}_{it} u_{it}, \tag{11}$$

where $\xi' = (1, \mathbf{0}_{17})$ is a vector selecting τ from the vector of all regression coefficients, and

[23] 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. For more details see Athey and Imbens (2016).

u_{it} is the residual term from the model estimated on $\mathcal{P}$.[24]

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

$$\rho_{it} = \xi' \left(\frac{1}{N_{\mathcal{P}}} \sum_{(j,s)\in\mathcal{P}} \mathbf{x}_{js}\mathbf{x}'_{js} \right)^{-1} \cdot \mathbf{x}_{it} u_{it}. \tag{12}$$

Using the CART algorithm by Breiman *et al.* (1984) on transformed outcomes (12), 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.*:

$$\frac{1}{N_{C_1}} \left(\sum_{(i,t)\in C_1} \rho_{it} \right)^2 + \frac{1}{N_{C_2}} \left(\sum_{(i,t)\in C_2} \rho_{it} \right)^2, \tag{13}$$

which, as Athey *et al.* (2019) show for a more general case, is consistent with maximizing criterion (8). 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:

$$\mathrm{Var}(\hat{\tau}(\mathbf{z})) = \xi' V(\mathbf{z})^{-1} H(\mathbf{z}) \left(V(\mathbf{z})^{-1} \right)' \xi. \tag{14}$$

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 (14) 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

[24] 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 (6), x_n is $\tau_{\mathcal{P}}$ and x_{n+1} corresponds to τ_{C_k} in the textbook Newton-Raphson root-finding algorithm.

x	-125,000	-100,000	-75,000	-50,000	-25,000	0	25,000
$P(\hat{\tau}_i^{\text{🌳}}<x)$	0.051	0.180	0.363	0.619	0.836	0.909	0.967
$P(\hat{\tau}_i^{\text{🌳}}+1.96{*}se<x)$	0.015	0.075	0.234	0.465	0.706	0.861	0.905
$P(\hat{\tau}_i^{\text{🌳}}-1.96{*}se>x)$	0.860	0.683	0.453	0.256	0.093	0.061	0.005

Table 3: Distribution of cumulative individual earnings losses, estimates from a generalized random forest. The first row shows the empirical cumulative distribution of earnings loss estimates $\hat{\tau}_i^{\text{🌳}}$. Row 2 reports the fraction of workers with point estimates statistically significantly below x, whereas row 3 shows the fraction of workers with point estimates statistically significantly above x (at the 95 % confidence level).

$\hat{\tau}(\mathbf{z})$, 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 (14). Thanks to this, our standard errors measure estimation accuracy affected by both machine-learning uncertainty and estimation noise.[25]

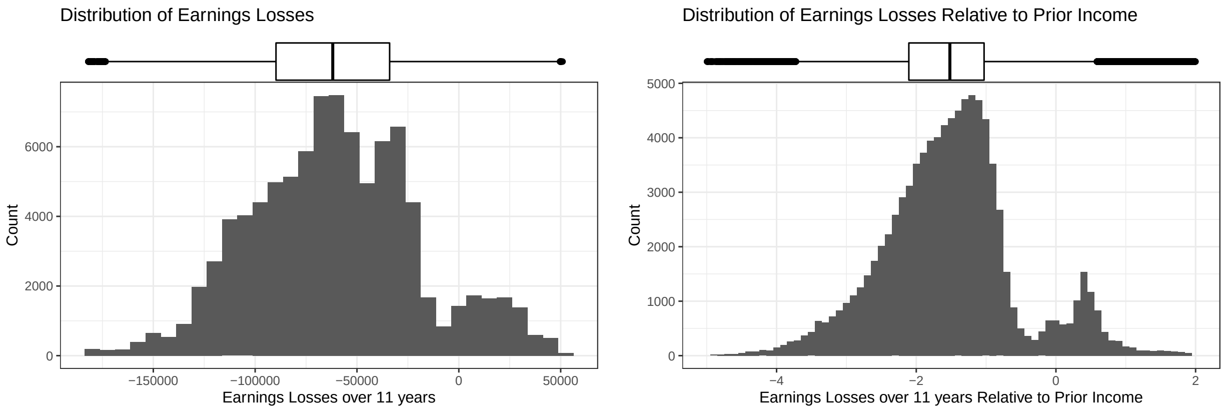
V. Heterogeneous Scarring Effects of Job Loss

In this section we use the results from the machine learning approach to answer two questions. On the one hand, we are interested in the heterogeneity behind the average costs of job displacement. Second, we leverage that we also know how the heterogeneity is related to observable worker and job characteristics. This allows us to disentangle the sources of earnings losses. In the analysis, we grow a forest consisting of 10,000 trees, with the minimum leaf size equal to 1,600 person-year observations. In each considered split we draw 6 randomly chosen partitioning variables. Observations are clustered at the worker level.

A. Heterogeneity in Earnings Losses

The generalized random forest provides an estimated treatment effect for each individual worker based on his characteristics. We are thus in a position to document how the scarring effect of job displacement differs across workers. We start with plotting the distribution of cumulative earnings losses τ_i over 11 years after the displacement event. Figure 3a and Table 3 shows that the average effect hides an enormous amount of heterogeneity among workers. Median cumulative earnings losses are €62,900, or 1.52 times the yearly annual

[25] An example illustrating how the algorithm works on simulated data is relegated to the Online Supplement.



(a) Absolute Losses (in Euros) (b) Relative to prior income

Figure 3: Distribution of cumulative earnings losses over the first 11 years after job displacement. Estimates from a generalized random forest

median earnings in the year before the displacement event. But 25% of individuals face staggering earnings losses larger than €89,763, whereas around 10% of individuals experience no earnings losses. Table 3 also shows that this heterogeneity is not generated through noisy point estimates of individual earnings losses. For 7.5 % of the population we can rule out (at the 95 percent confidence level) earnings losses smaller than €100,000, whereas for 6% of workers we can rule out any detrimental effects on earnings.

Another angle of heterogeneity that we document is the distribution of earnings losses relative to income before displacement. It is noteworthy that in the presence of dispersion in pre-displacement income, heterogeneous scarring effects in absolute terms does not translate automatically to heterogeneity relative to prior income.[26] However, in our study we find that individuals with higher absolute losses also suffer higher relative losses. This is shown by the high correlation between both measures of 0.7. Row 1 of Table 4 shows that also the relative heterogeneity is substantial, with interquartile range amounting to more than one. In addition to this, changes in log wages measure percentage changes relative to pre-displacement wages. The conclusions drawn from the random forest grown to maximize heterogeneity in losses in log wages are essentially unchanged.

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

[26]For instance, if every displaced worker lost exactly 10% of the income, then still there would be heterogeneity in losses in absolute terms but not heterogeneity in relative terms at all.

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 4: 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).

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 4 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.

To study difference in the evolution of earnings losses, as well as employment and wage losses we proceed by binning workers into quartiles according to their estimated earnings losses. Figure 4 plots the evolution of earnings losses, employment losses and log wage losses by estimating the Event-study specification from equation (1) separately for every quartile of estimated earnings losses, alongside with the conventional 95% confidence intervals. The figure reconfirms that workers face significantly different scarring effects of job loss. First of all, it is clearly visible that the heterogeneity does not only originate from different short term evolution of earnings, but are persistent throughout the 10 year window after job displacement. The group of workers with the lowest estimated losses (Q4) gain in terms of yearly income from job displacement in the long run. They only face short employment losses of about one month in the year of displacement, but recover the employment losses almost completely afterwards. The income gains result from an increase in wages by 3-4 percent. This is in stark contrast to the other three quarters of workers, who all face long-term earnings losses. The quartile of workers with the highest estimated earnings losses (Q1) experience employment losses of one and a half years and conditional on reemployment, depressed wages by 20%. Workers with high absolute earnings losses therefore also face high wage losses in relative terms. This group of workers also face significant employment losses. Even 10 years after separation, days employed is still depressed by over one month, while employment losses for all other groups recovered almost fully. Visually, it appears that heterogeneity in earnings losses arise mostly through differences in wage changes. Next, to address this question more precisely, we formally decompose earnings losses into employment

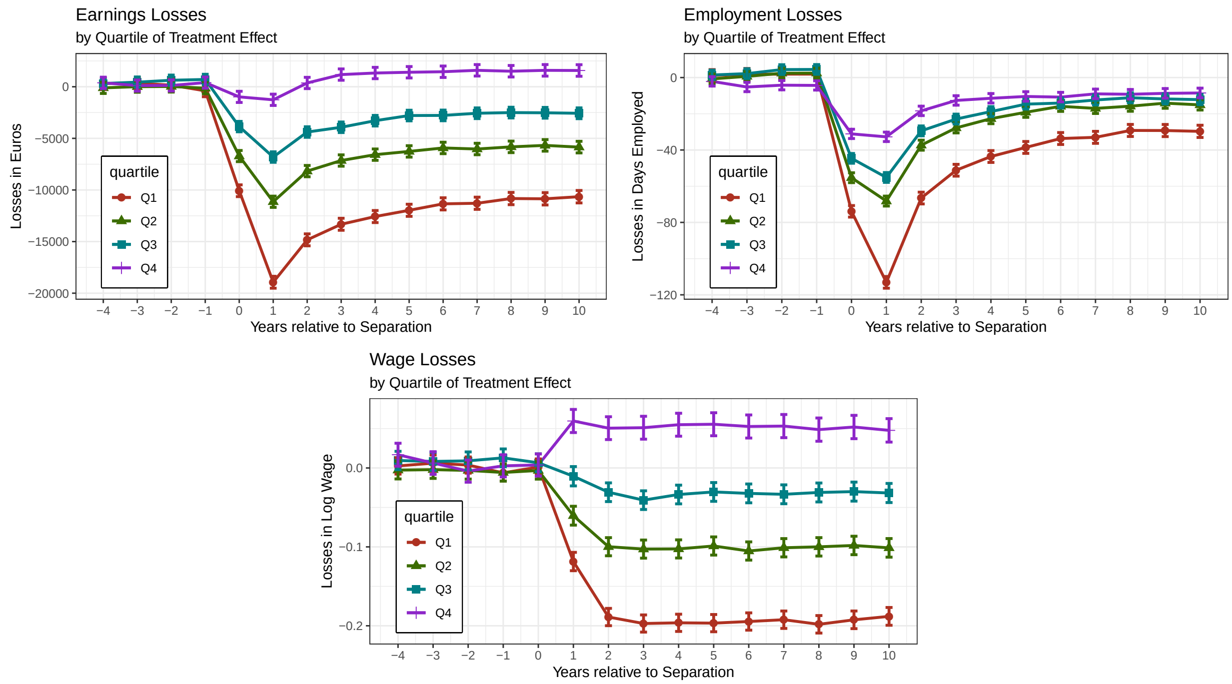


Figure 4: Earnings, employment and wage losses of displaced workers - Eventstudy regression estimates, by quartile of estimated earnings losses, estimates from a random forest. Period 0 corresponds to the separation year. Earnings and days employed are computed for the whole year, log wages are computed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

and wage losses.

B. 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. To this end, we grow two additional forests maximizing heterogeneity in the treatment effect τ with changing the dependent variable y in equation (3) to log-wages and employment. Figure 5 depicts the joint distribution of the estimated losses in wages in log points and employment in days through 11 years after the job displacement. Both estimates are positively correlated with each other with the correlation coefficient amounting to 0.475. The wage of displaced workers after the layoff event is on average 7% lower than their counterparts from the control group. That said, 19% of workers do not suffer from any wage losses, but gain in terms of

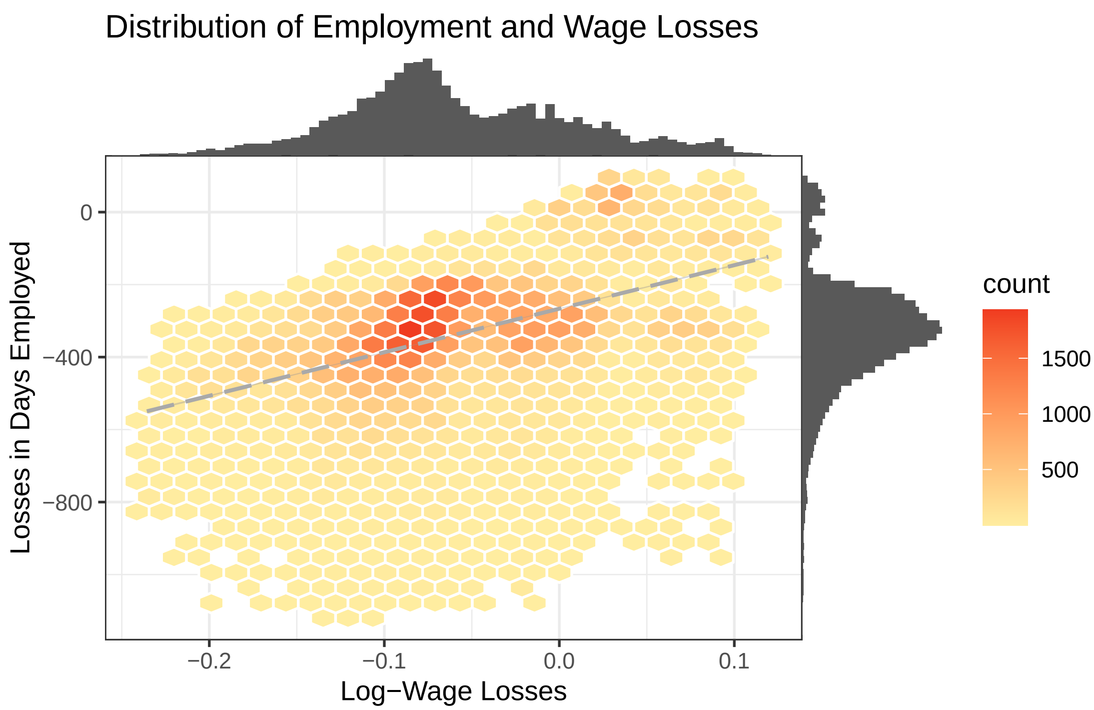


Figure 5: Joint distribution of wage and employment losses estimated by two independent random forests. Marginal Distributions are at the edges of the plot.

wages. The average employment loss over 11 years after the displacement is equal to 336 days and for 90% workers the loss is longer than half a year.

More 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.* (2020), 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:

$$\begin{aligned}
\mathbb{E}[\Delta] &= \mathbb{E}[y^C] - \mathbb{E}[y^D] = \mathbb{E}[N_d^C w^C] - \mathbb{E}[N_d^D w^D] \\
&= \mathbb{E}[N_d^C]\mathbb{E}[w^C] - \mathbb{E}[N_d^D]\mathbb{E}[w^D] + \mathrm{Cov}(N_d^C, w^C) - \mathrm{Cov}(N_d^D, w^D) \\
&= \left(\mathbb{E}[N_d^C] - \mathbb{E}[N_d^D]\right)\mathbb{E}[w^C] + \mathbb{E}[N_d^D]\left(\mathbb{E}[w^C] - \mathbb{E}[w^D]\right) + \Delta\mathrm{Cov}(N_d, w) \\
&= \Delta\mathbb{E}[N_d]\mathbb{E}[w^C] + \mathbb{E}[N_d^D]\Delta\mathbb{E}[w] + \Delta\mathrm{Cov}(N_d, w). \qquad (15)
\end{aligned}$$

Figure 6 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

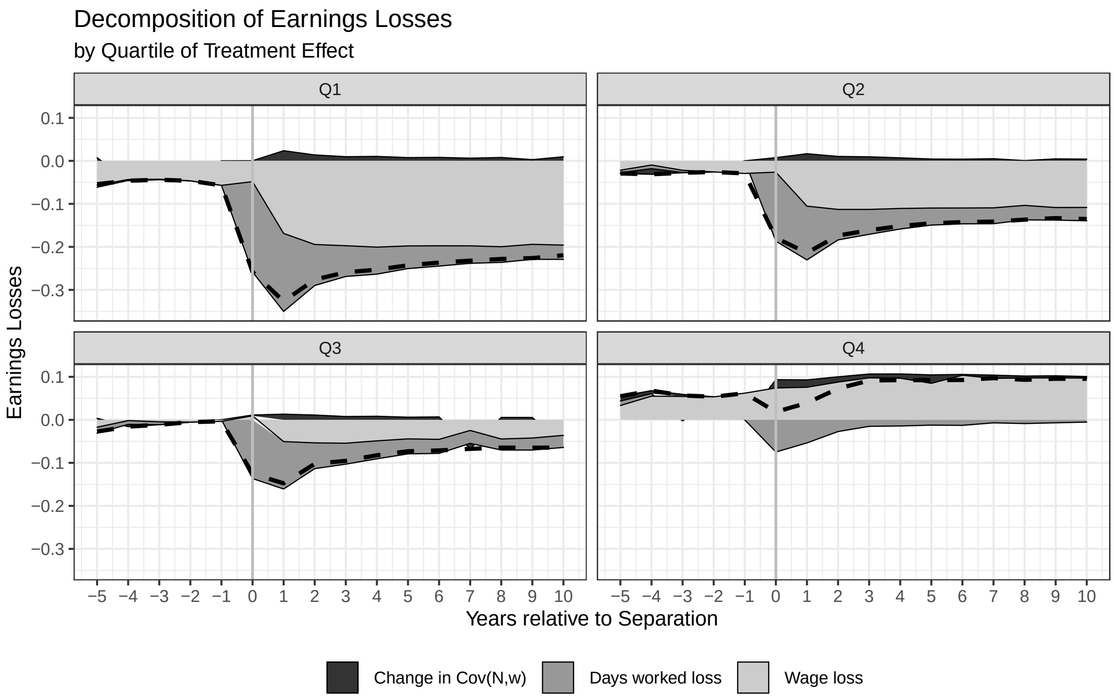


Figure 6: Decomposition of earnings losses by quartile of predicted treatment effect using equation (15). Estimates from a generalized random forest. Broken line indicates total earnings losses.

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.

C. 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 5 reports descriptive statistics broken down by quartile of estimated earnings losses. The workers with the highest earnings losses have above average tenure and income, are employed at better paying firms, and are

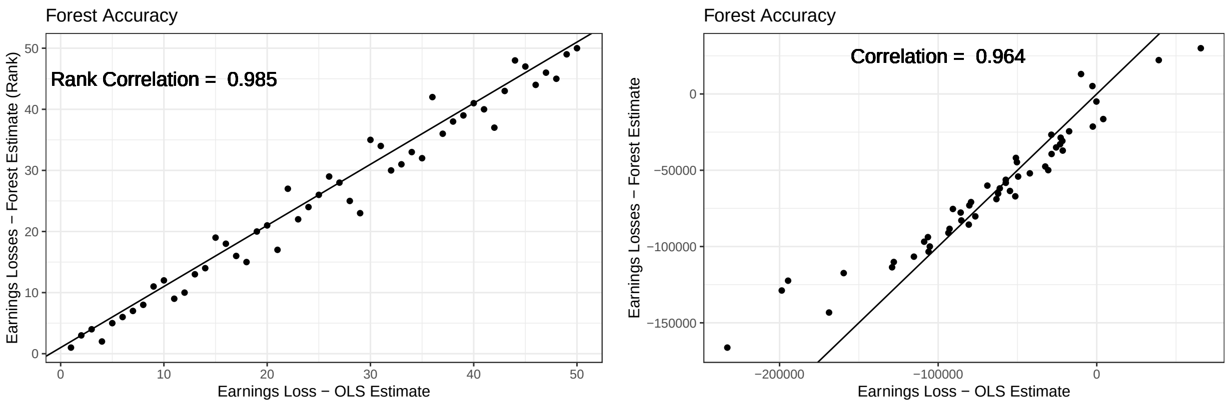
	Q1	Q2	Q3	Q4
Blue Collar	0.482	0.511	0.588	0.509
Manufacturing	0.642	0.670	0.540	0.242
Austrian	0.786	0.795	0.774	0.801
Firm FE	8.641	7.323	5.655	3.099
Firm Size	6.744	6.182	6.070	6.969
Match Quality	6.151	6.027	5.249	4.808
Job Tenure	7.898	7.351	7.278	7.295
Herfindahl Index	5.343	4.755	5.366	6.414
Avg. Firm FE	5.369	5.276	6.053	6.116
Industry UE-Rate	5.415	6.378	5.929	4.473
Regional UE-Rate	5.945	5.575	5.564	6.082
Recession	0.403	0.365	0.318	0.234
Number of Firms	2.113	2.364	2.455	2.620
Age	41.161	37.796	36.136	36.190
Firm Sepa. Rate	4.897	5.440	5.663	5.608

Table 5: Mean baseline characteristics for each quartile of estimated treatment effects. Estimates from a generalized random forest

more likely to work in the manufacturing sector, have a white collar occupation, and have worked for fewer firms over their careers. 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 5 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 17), so it is hard to draw definite conclusions from these compositional differences. Before this question will be tackled in section VI, we evaluate how accurately our random forest estimates the heterogeneity of losses.

D. 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



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

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

Figure 7: 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

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 (1) using OLS. We then compare the OLS earnings loss estimates with the results of our random forest. The left panel of Figure 7 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 0.985. The right panel plots the OLS estimates against the earnings loss estimates from the random forest. The correlation is with 0.964 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.

VI. Sources of Earnings Losses

The previous section documented the heterogeneous scarring effects of job displacement. The results were already indicative of which channels are important to explain earnings losses. Workers with the highest earnings losses have higher firm fixed effects, higher match effects and higher job tenure for example. But this information is still not enough to identify which factors are causing higher earnings losses. First, it just documents how workers with higher earnings losses differ from workers with little losses and not which factor is leading to higher losses. Second, as workers with high losses differ in many characteristics from workers with little earnings losses, it is hard to single out the most important factors. One succinct way to understand which factors are the most important is the variable importance measure obtained from the random forest, to which we turn next.

A. Which channels are the most important?

A compact way to assess which factors are the most important in explaining differences in earnings losses is the occurrence frequency of variables in the splitting criteria. Variables chosen more frequently have a higher contribution in explaining the heterogeneity of scarring effects. The firm-fixed effect was by far the most often chosen by the algorithm and has been used in 38% of all splits of the depth level lower than 4. That said, this raw statistics might be misleading. Due to a top-down greedy approach, variables chosen first tend to be more important. The importance of variables chosen earlier is underestimated.[27] For this reason, it is a common practice to compute a depth-adjusted variable frequency, which puts higher weights for splits selected first. Figure 8 shows the adjusted frequency with a decay exponent equal to -2, which is the default value in most of the machine learning literature.[28] For this measure, the firm fixed effect is even more important relatively. This result is quite striking. Due to variable randomization, random forests tend to spread the importance out across many variables.[29] However, in our exercise, the firm fixed effect continues being the dominating variable. The regional average firm fixed effect is the second most important variable and the firm size is the third one (with almost a tie with the sectoral unemployment rate).

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 linearly project the three most important variables, *viz.:* firm wage premia ψ_i, regional average firm wage premia ψ_i^R, and firm size $fsize_i$, onto the individual cost of

[27] To understand this property, imagine two binary partitioning variables, z_1 and z_2, where z_1 is more important than z_2. The tree-building algorithm will always choose z_1 first and z_2 will be chosen later and will be conditioned on two values of z_1. As a result, the more important variable z_1 will occur in the splitting criteria only once, while a variable of the second order, z_2, will be used twice.

[28] It means that split frequencies for nodes of depth k is 50% less important than split frequencies for nodes of depth $k-1$.

[29] For more details see Efron and Hastie (2016, pp. 331–332).

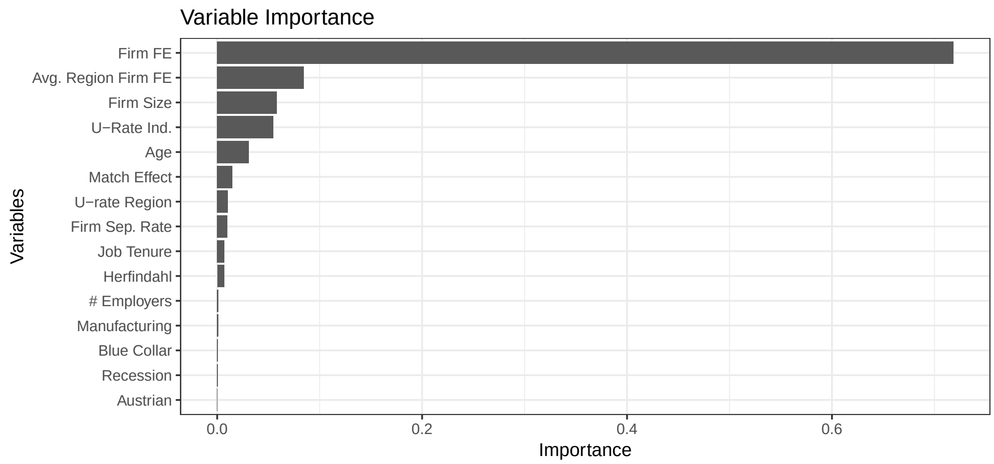


Figure 8: Depth-adjusted variable frequency in splits in the GRF with a decay exponent equal to -2 and the maximum depth level of nodes equal to 4. All values sum to 1.

job loss $\hat{\tau}_i^{🌳}$ estimated using our random forest and perform a variance decomposition. We consider the variance decomposition of both earnings losses in terms of total income, as well as log-wage losses, which measure percentage losses in terms of pre-displacement wages. In practise, we estimate the following model:

$$\hat{\tau}_i^{🌳} = \beta_0 + \beta_1 \psi_i + \beta_2 \psi_i^R + \beta_3 fsize_i + \epsilon_i. \tag{16}$$

Table 6 depicts the variance decomposition of the estimated model. Here again, the firm wage premia have the overwhelming contribution independently whether we consider absolute earnings losses, or relative log-wage losses. Its variance alone accounts for almost 63% of variability in individual earnings losses (slightly reduced by 5.4 percentage points due to interactions terms) and a staggering 73% of the variation in individual log-wage losses. Both measures of variable importance indicate that lossses in firm wage premia is the most important factor in explaining earnings losses. Nevertheless, in our analysis we are able to go even further and we investigate the counterfactual changes in the value of one variable while keeping all others at their empirical level, which we present next.

B. Conditional Average Treatment Effects

How do earnings losses vary by changing one factor at a time? Do these variables affect earnings losses through declines in wages, or through losses in days employed, or both? To answer this question, we take advantage of our generalized random forests where we

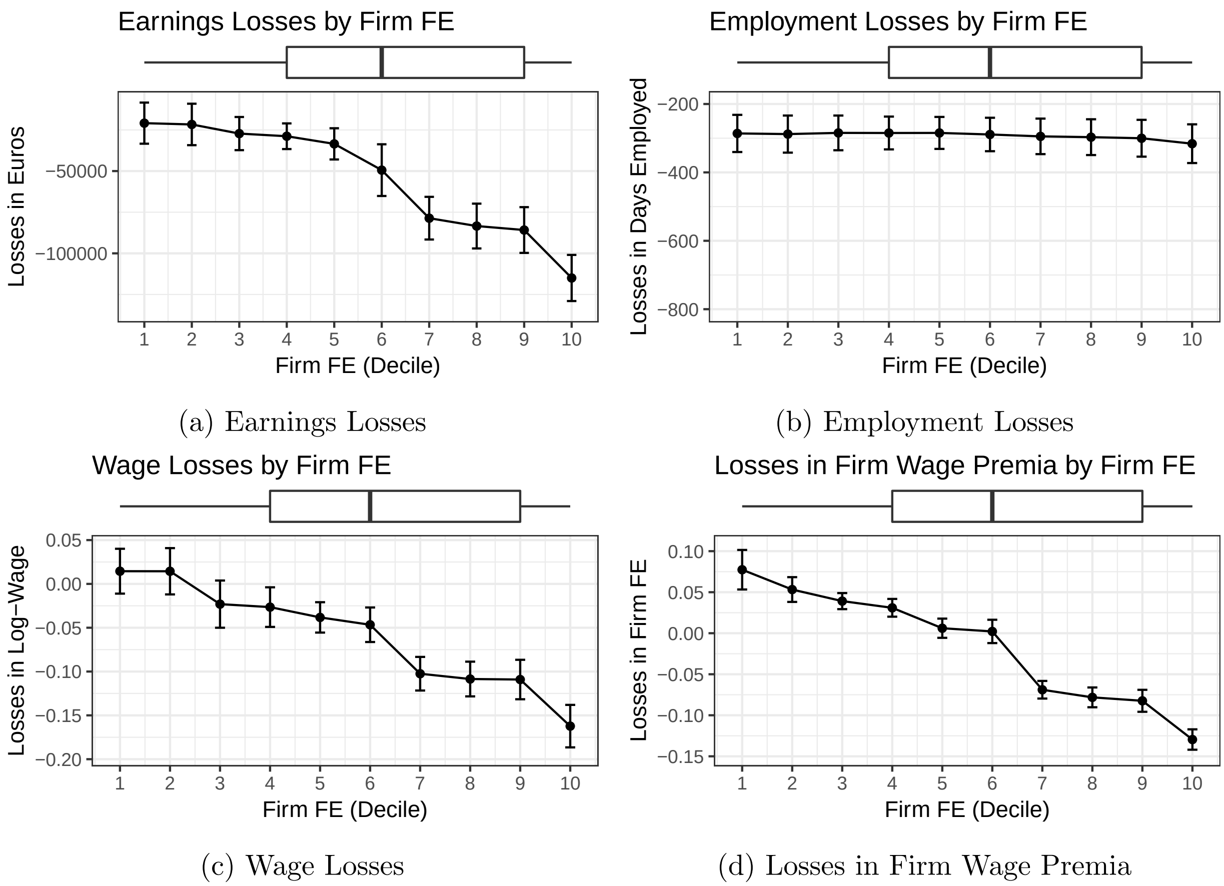
	Share of total variance	
	Earnings Losses	Log-wage Losses
Var($\hat{\tau}_i^{\circledast}$)	1.000	1.000
Var($\beta_1\psi_i$)	0.627	0.728
Var($\beta_2\psi_i^R$)	0.065	0.037
Var($\beta_3 fsize_i$)	0.002	0.001
2Cov($\beta_1\psi_i, \beta_2\psi_i^R$)	-0.054	-0.044
2Cov($\beta_1\psi_i, \beta_3 fsize_i$)	0.004	0.004
2Cov($\beta_2\psi_i^R, \beta_3 fsize_i$)	0.001	0.001
Var(ϵ_i)	0.355	0.274

Table 6: Variance decomposition for cost of job loss in terms of yearly earnings and log-wages. Calculations based on regression model (16)

can compute losses $\tau(\mathbf{z})$ conditional on particular realizations of the partitioning variables, $\mathbf{z}$. We compute how losses in earnings, employment days, log-wage and firm wage premia vary, by changing the realization of one factor at a time, while holding all other variables fixed at their median. This way we measure the impact of changing one channel at a time, while holding confounding factors constant. 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.

Figures 18 and 19 in the appendix report how earnings losses change with the 15 different partitioning variables considered, sorted by the variable importance measure. We categorize all continuous variables in deciles according the the overall distribution of workers employed on the reference day. Since displaced workers and the selected control group might differ from the general population, we also include a boxplot of the distribution of our sample over the variable of interest on top of every plot. Alongside the point estimate, the figures also contain 95% confidence intervals that account for the uncertainty arising from both, the machine learning procedure and the estimation procedure.[30] We scale our results for earnings losses and employment losses by 11, so that the results can be interpreted as 11-year cumulative losses. To enhance comparability, we are holding the y-scale constant across figures. This way it is easy to see which variables are more important drivers of earnings losses. Moreover, as an additional analysis, the same exercise has been carried out on the random forest grown to maximize the heterogeneity in changes in log wages, which measure percentage losses

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



(a) Earnings Losses

(b) Employment Losses

(c) Wage Losses

(d) Losses in Firm Wage Premia

Figure 9: 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

in terms of pre-displacement wages. The conclusions on the variable importance remain essentially the same (see Figures 22, 23, and 24 in Appendix D). This is not surprising, given the high correlation between absolute and relative losses documented before.

B.1. Firm Wage Premia

Figures 9 and 18 confirm the finding 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 €20,873, whereas workers formerly employed in the top decile paying jobs forgo more than 5 times as much, with earnings losses amounting to almost €114,969. These heterogeneous effects are

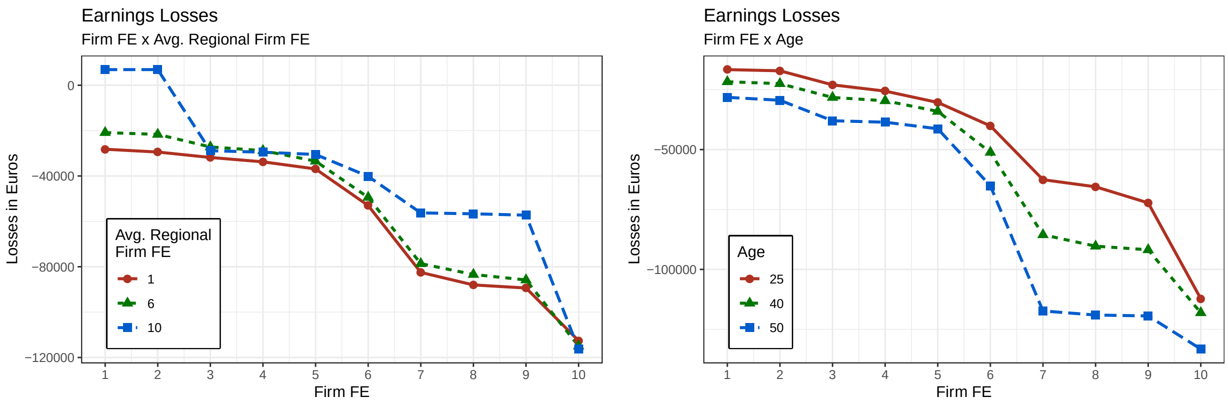


Figure 10: GRF estimates of the cumulative 11-year earnings losses for different values of the firm fixed effect and the age. All other variables are set to their median values.

also very precisely estimated. The confidence intervals are not wider than ±€10,000.

To understand whether this 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 more than 16% for workers at the highest paying firms. Panel (d) 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-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. Moving a worker with median characteristics from a region in the bottom decile of the average firm pay distribution to the highest reduces the estimated earnings losses by almost €13,000. Interestingly, compared to other variables, this slope seems low, given that it is the second most important factor in explaining the heterogeneity in earnings losses. This

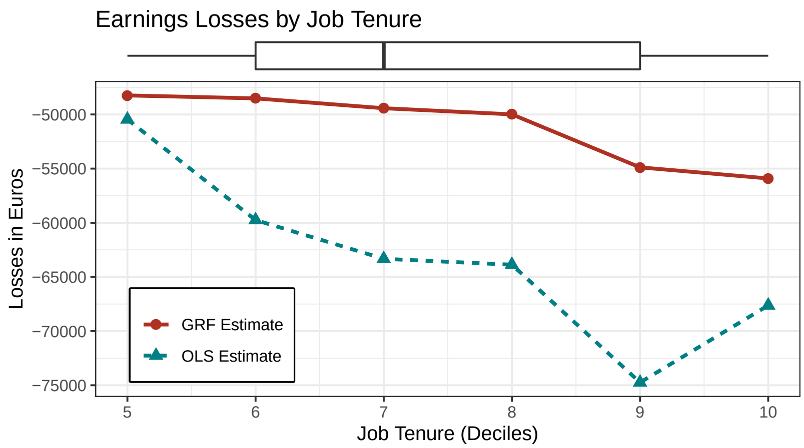


Figure 11: Comparison between Sample Splitting and GRF for Earnings Losses by Job Tenure.

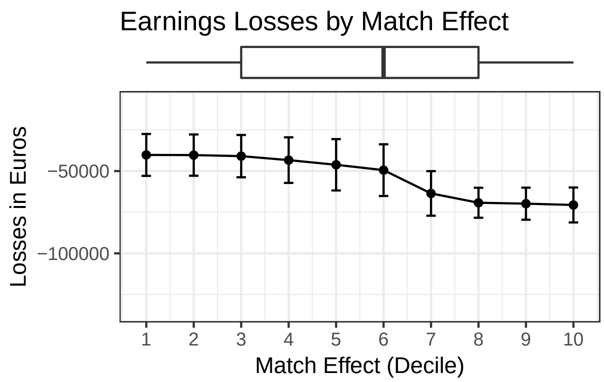
suggests that there are strong interactions effects with other variables, which is confirmed in the left panel of Figure 10. It reports the effect of firm fixed effects by the average regional firm pay premium. Workers employed in low fixed effect firms, but located in high firm wage premium regions do not experience any income losses.

We also consider a robustness exercise, where we compute the firm fixed effects on observations before 1998, and re-estimate the random forest on earnings loss observations after 1998. This rules out that the mass layoff event affected our measure of earnings losses. Appendix E shows that the main findings are essentially unchanged.

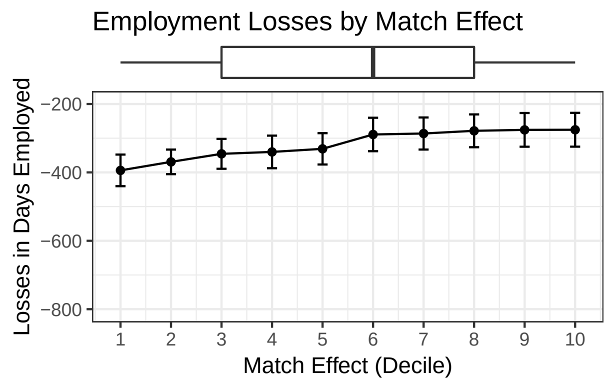
B.2. Worker's Job Tenure

Perhaps the most prominent theory about the sources of earnings losses is that workers lose job specific human capital. Several papers have shown that workers with higher tenure experience higher losses, e.g. (Topel, 1990; Jacobson *et al.*, 1993). This is in stark contrast to our findings. Figure 11 shows that earnings losses only marginally increase from €48,000 for a low tenure workers to €55,000 for high tenure, but otherwise observationally identical worker. Figure 11 compares the estimates from our forest and to sample splitting, where we re-estimate earnings losses using OLS for each tenure group. A traditional sub-group analysis would yield a much higher slope in tenure.[31] This comparison shows that it is crucial to hold confounding factors constant, as high tenure workers are different in many

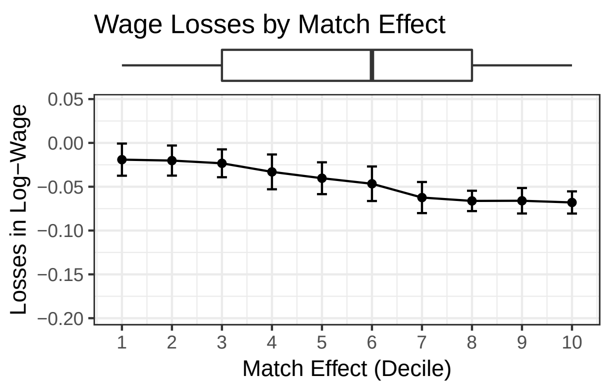
[31] The overall level difference comes from the fact that the average worker has slightly higher earnings losses than a worker with median characteristics.



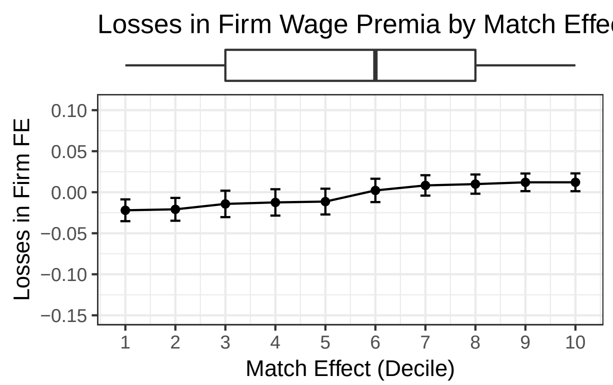
(a) Earnings Losses



(b) Employment Losses



(c) Wage Losses



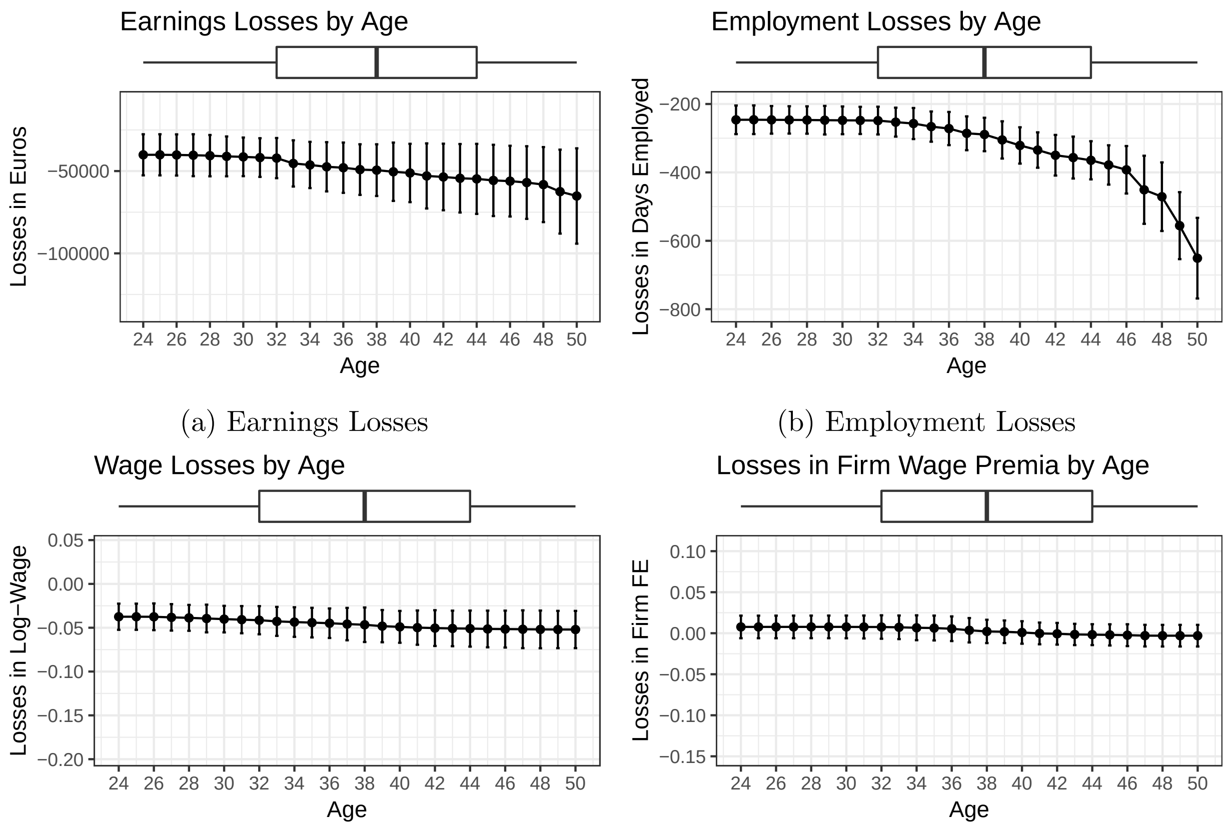
(d) Losses in Firm Wage Premia

Figure 12: 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

other characteristics from low tenure workers.

B.3. Job Match

In contrast to losses in firm specific human capital, we find evidence that part of the earnings losses arise from losing particularly good matches. Figure 12 displays how losses in earnings, employment, wage and firm premia are affected by the match specific wage component, which is captured by the residual term of equation (5). Panel (a) shows that high match effect workers face higher earnings losses compared to observationally similar workers with low levels of the match effect. Although workers in high match quality jobs face lower employment losses, they face higher wage losses. The shape of wage losses closely mirrors the shape of earnings losses. Panel (d) shows that varying income while holding other



(a) Earnings Losses

(b) Employment Losses

(c) Wage Losses

(d) Losses in Firm Wage Premia

Figure 13: 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

factors constant does not affect losses in firm wage premia. This is evidence that workers in particularly good matches fall off the match quality ladder as in Jovanovic (1979).

B.4. Age

Workers' age also plays an important role in understanding earnings losses. Workers aged 30 and below face losses of about €40,000, whereas otherwise identical older worker tend to experience higher losses. Figure 13 shows that age affects earnings losses mostly through employment losses. Workers aged 50 experience cumulative employment losses of 22 months, compared with less than 8 months for young workes. Wage losses are not influenced by worker's age. The right panel of Figure 10 reveals an interesting interaction between worker's age and firm's wage premia. Worker's age amplifies the impact of firm's wage premia. Job

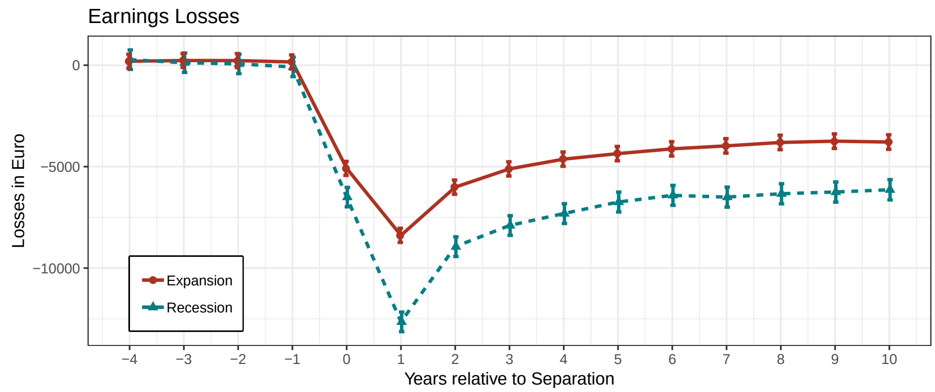


Figure 14: Earnings losses in recessions and expansions using the Event-study specification (1). Observations were split into recessions and expansions categorized according to OECD definitions.

losses at high paying firms are significantly more consequential for older workers. While workers aged 25 displaced from firms with wage premia between the seventh and ninth deciles face losses from €62,000 to €72,000, the losses of workers aged 50 are almost twice as high, amounting to around €120,000. Panel (d) of Figure 13 shows that reemployed older workers in fact do not face higher losses in firm's wage premia. This suggests that employment losses are more pronounced for older workers separating from well paying jobs.

B.5. Cyclicality of Earnings Losses

Three variables speak to the cyclicality of earnings losses in our study. We study how losses change with whether the Austrian economy is in a recession according to OECD definition, as well as with the unemployment rate in the local labor market and the prior industry of the worker. As can be seen in Figure 19, earnings losses do not vary much with all business cycle indicators, when holding confounding factors constant. For example, a displaced worker faces only €1,400 higher 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.[32] In contrast, if we split our sample and re-estimate earnings losses using the Event-Study specification (1) separately for workers separating in recessions and expansions we find significantly higher

[32]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.

losses in recessions, as can be seen in Figure 14. This finding is consistent with similar observations documented by Davis and Von Wachter (2011) and Schmieder *et al.* (2020) 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 14: 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. Formally, we can disentangle both effects using the following decomposition:

$$\begin{aligned}\int \tau(\mathbf{z}|rec=1)\mathrm{d}F(\mathbf{z}|rec=1) - \int \tau(\mathbf{z}|rec=0)\mathrm{d}F(\mathbf{z}|rec=0) \quad &= \\ \underbrace{\int \left[\tau(\mathbf{z}|rec=1) - \tau(\mathbf{z}|rec=0)\right] \mathrm{d}F(\mathbf{z}|rec=1)}_{\text{Recession effect}} \quad &+ \\ \underbrace{\int \tau(\mathbf{z}|rec=0)\mathrm{d}F(\mathbf{z}|rec=1) - \int \tau(\mathbf{z}|rec=0)\mathrm{d}F(\mathbf{z}|rec=0)}_{\text{Compositional difference}}, & \end{aligned} \tag{17}$$

where $F(\mathbf{z}|rec=r)$ denotes the distribution of worker and job characteristics of displaced workers, $\mathbf{z}$, conditioned on the aggregate state of the economy, where rec is equal to 1 during a recession and 0 otherwise. Then $\int \tau(\mathbf{z}|rec=r)\mathrm{d}F(\mathbf{z}|rec=r)$ is the average treatment effect of job loss for all workers displaced conditioned on the aggregate state $rec=r$. In our application those 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 are in the position to identify the counterfactual treatment effect of job loss in good times for workers displaced in bad times. The average difference between a actual treatment effect and the counterfactual effect conditioned on no recession, $rec=0$, is the pure effect of a recession, while the remaining part of the right-hand side of (17) comes from compositional differences. While a traditional sample-splitting approach captures the overall effect, the left-hand side of (17), our random forest allows us to identify those two components, the recession effect and the effect of the compositional differences separately. Table 7 shows that the recession impact by itself is very small and accounts for merely 6% of the overall differences in earnings losses and the remaining part stems from compositional changes of displaced workers. A similar conclusion can be drawn from comparative statics with respect to the recession dummy while keeping all other variable at the average level as can be seen in bottom right panels of Figures 19 and 21.

Table 7

	Euros	Share
Total Difference	-19, 399.03	1
Recession Effect	-1, 155.95	0.06
Composition Effect	-18, 243.09	0.94

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

B.6. Other Factors

From all the remaining factors, firm size is the only variable that affects earnings losses somewhat. Losses slightly increase from €48,500 at the smallest firms in our sample to €62,500 for the largest firms. All other factors do not seem to play an important role, varying these variables while holding other factors constant barely changes the estimated earnings losses. The regional labor market concentration, indicators for blue collar, Austrian and manufacturing workers do not affect earnings losses. Further, the firm level job security prior to separation has no impact on earnings losses. Since displaced workers by definition lose their jobs, this factor operates entirely trough differential earnings evolution of the control group. Control group workers employed at less stable firms have indeed a significantly higher separation rate, but the results suggest that the non-masslayoff separations of the control group do not entail large earnings losses.[33]

C. Partial Dependence Plots

One potential criticism of the comparative statics in subsection B is that an individual, with median characterstics might not be representative for the whole population or even may not exist at all. Moreover, the focus in the analysis on such an artificial unit might generate misleading results due to plausible non-linearities and interactions between the partitioning variables which occur far from the median. The computed measures may be relevant only

[33] The control group's probability of separating from the original employer within 5 years increases from 15 percent to 44 percent over the range of our firm separation rate measure.

for the neighborhood of the median and may disregard other strong effects in the whole sample. We tackle this concern in two ways. First, we created a companion web-applet (`https://gulyas-pytka.app/earnloss`) where the reader can interactively explore how earnings losses change with all the variables considered in this study. Thus, earnings losses can be estimated for any worker or job characteristic very easily.[34] Second, we use partial dependence plots proposed by Friedman (2001) to better understand how *on average* a single variable affects 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:

$$\mathbb{E}_{\mathbf{z}_{-k}} \hat{\tau}(z_k = \overline{z}; \mathbf{z}_{-k}) = \int \hat{\tau}(z^k = \overline{z}; \mathbf{z}^{-k}) \mathrm{d}F(\mathbf{z}^{-k}), \tag{18}$$

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 20 and 21 in the appendix depict partial dependence plots 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.

VII. Policy Targeting

Given the large amount of heterogeneity in earnings losses we document that any policy intervention should perhaps be targeted. In this section we present an example of how our study can support such decisions. To this end, we employ policy targeting with a policy tree in the spirit similar to Athey and Wager (2017). Suppose that the government wants to implement a policy intervention aimed at individuals suffering earnings losses above the median level during mass layoffs. At the same time the decision rules should be simple and rely on criteria understood by everyone.[36] Therefore, using individual earning losses

[34] Due to the computational burden, the web applet contains coarser grids and the forest consists of 2,000 trees instead of 10,000. Therefore, the results can slightly differ.

[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.

[36] In the EU countries, policy makers are obliged to provide grounds of their decision. In particular, the General Data Protection Regulation (EU) 2016/679 introduced by European Parliament and Council of European Union (2016) gave "the right to explanation" to all individuals that are subject to decisions

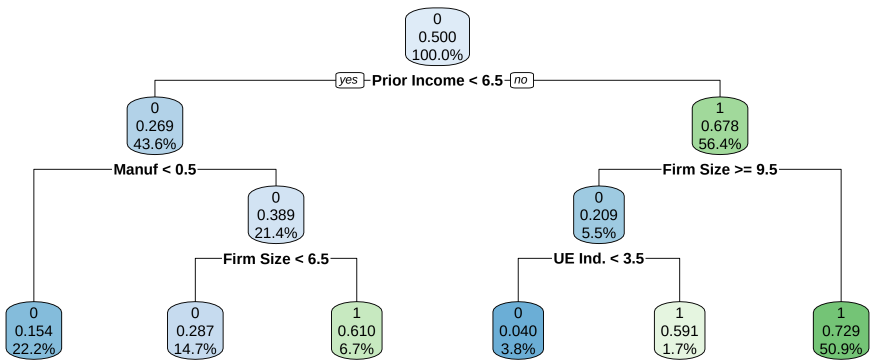


Figure 15: Classification tree predicting earnings losses above the median level. On the top there is the most common value. In the middle there is a fraction of observations with earnings losses above the median. In the bottom there is a fraction of observations in the global sample.

identified through our generalized random forest, we build a classification tree. In this tree we predict whether an individual exhibits earnings losses above the median level ($y = 1$ if so, $y = 0$ otherwise) using all partitioning variables except those that might be difficult to measure at the time of the lay-off, *i.e.* firm fixed effects (at individual and regional level), match effects, and labor-market concentration. Instead we included income prior to displacement as an additional variable that can easily be observed. For simplicity the max depth of the tree was set to 3.

Figure 15 presents the classification tree grown using the CART algorithm by Breiman *et al.* (1984). As can be seen, there are three subgroups suffering larger earnings losses. The first group is constituted by high-income individuals hired at all companies but the largest ones. In this group 72.9% are characterized with high losses. The second group consists of high-income individuals hired at the largest firms displaced in times when unemployment of their sector is not at the lowest level. In this group 59.1% suffer higher earnings losses. The last group are low-income individuals working for large non-manufacturing companies, where 61% suffer higher earnings losses. The targeting performance of the decision rule coming from the tree is presented in Table 8.[37] The targeting policy using this very

made by automated decision-making algorithms. Article 12 states *"The controller shall take appropriate measures to provide any information (...) relating to processing to the data subject in a concise, transparent, intelligible and easily accessible form, using clear and plain language, in particular for any information addressed specifically to a child"* and Article 13 gives the right to *"meaningful information about the logic involved."*

[37]The CART algorithm provides probability of having high losses, $P(\mathbf{z})$. In the prediction stage, we chose

Prediction	*Actual Treatment Effect* Low	High	Sum
Low	32.90%	7.79%	40.69%
High	17.10%	42.21%	59.31%
Sum	50.00%	50.00%	100.00%

Table 8: Confusion matrix for the classification tree from Figure 15.

simple model has a surprisingly high accuracy, 75.11% observations are classified correctly.[38] This implies a substantial lift over a policy without targeting and with a completely random intervention assignment, for which the accuracy is equal to 50%.[39] The discussed example illustrates that even very simple rules obtained from our study can provide a meaningful insight into designing policies with higher accuracy.

VIII. Conclusions

We implement a generalized random forest (Athey *et al.*, 2019) to a difference-in-difference 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 2017, we show that there is substantial heterogeneity in earnings losses across individuals. While the median cumulative earnings losses over 11 years after job displacement amount to €62,000, almost 1/5 face earnings losses greater than €100,000. Almost 10% of workers gain in terms of earnings. Second, we are able to predict earnings losses for workers that fail the usual sample restrictions about tenure and firm size typically applied in the literature. We show that overall these sample restrictions do not select individuals with particularly different earnings losses. Third, the machine learning procedure allows us to conduct a horse race between many competing theories about earnings losses, while controlling for confounding factors. We find that the

$\hat{y} = \mathbb{1}_{P(\mathbf{z})>0.5}$. We motivate this decision with the fact that our sample is perfectly balanced thanks to the propensity score matching from Subsection II.C. Alternatively, we could regularize the threshold parameter to maximize the accuracy measure. Nonetheless, bearing in mind that the purpose of our exercise is to provide the simple decision rule, we do not do it.

[38] In comparison, a tree built with all variables (including those hard to measure) has accuracy of 85.14%.

[39] The other measure of the model performance are also high. The true positive and true negative rates are equal to 84.42% and 65.8%, respectively. The true positive rate is defined as the fraction of individuals with actual high treatment effects with the correct prediction and the true negative rate is the fraction of individuals with actual low treatment effects with the correct prediction.

pre-displacement firm wage premium is by far the most important factor. Holding all other channels fixed, the counterfactual earnings losses rise from €21,000 for workers employed at the lowest paying decile of firms to almost five times as much for workers displaced from the highest paying decile of firms. Further, the variation in firm wage premia alone explain close to 65% of the overall variation in earnings losses.

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 94% 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 each channel.

We further show how our machine learning estimates can be used to derive a simple decision rule for targeting high earnings loss individuals. Hence, we argue that the proposed theory-agnostic and completely data-driven random forest is especially well suited for this setup. Our overall findings provide evidence that earnings losses can be understood by mean reversion in firm rents and losses in match quality, rather than by a destruction of firm-specific human capital.

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. (2016). Recursive partitioning for heterogeneous causal effects. *Proceedings of the National Academy of Sciences*, **113** (27), 7353–7360.

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

— and Wager, S. (2017). Efficient policy learning. *arXiv preprint arXiv:1702.02896*.

Azar, J., Marinescu, I. and Steinbaum, M. I. (2017). *Labor Market Concentration*. Tech. rep., National Bureau of Economic Research.

Bagger, J. and Lentz, R. (2018). An empirical model of wage dispersion with sorting. *The Review of Economic Studies*, **86** (1), 153–190.

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.

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.

Carrington, W. J. and Fallick, B. (2017). Why do earnings fall with job displacement? *Industrial Relations: A Journal of Economy and Society*, **56** (4), 688–722.

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

Davis, J. and Heller, S. B. (2017). *Rethinking the benefits of youth employment programs: The heterogeneous effects of summer jobs*. Tech. rep., National Bureau of Economic Research.

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.

Eeckhout, J. and Kircher, P. (2011). Identifying sorting in theory. *The Review of Economic Studies*, p. rdq034.

Efron, B. and Hastie, T. (2016). *Computer Age Statistical Inference: Algorithms, Evidence, and Data Science*. New York, NY, USA: Cambridge University Press, 1st edn.

European Parliament and Council of European Union (2016). General data protection regulation. *Council (EU) 2016/679*, https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:32016R0679.

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. (2018). Spillover Effects of Mass Layoffs. *Journal of the European Economic Association*.

Gulyas, A. (2018). Firm dynamics and labor market sorting.

Guvenen, F., Karahan, F., Ozkan, S. and Song, J. (2017). Heterogeneous scarring effects of full-year nonemployment. *American Economic Review*, **107** (5), 369–73.

Hagedorn, M., Law, T. H. and Manovskii, I. (2017). Identifying equilibrium models of labor market sorting. *Econometrica*, **85** (1), 29–65.

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

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.

Ichino, A., Schwerdt, G., Winter-Ebmer, R. and Zweimüller, J. (2017). Too old to work, too young to retire? *The Journal of the Economics of Ageing*, **9**, 14–29.

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

Jarosch, G. (2015). Searching for job security and the consequences of job loss.

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

Jung, P. and Kuhn, M. (2018). Earnings losses and labor mobility over the life cycle. *Journal of the European Economic Association*, **17** (3), 678–724.

Knaus, M., Lechner, M. and Strittmatter, A. (2017). Heterogeneous employment effects of job search programmes: A machine learning approach.

Krolikowski, P. (2017). Job ladders and earnings of displaced workers. *American Economic Journal: Macroeconomics*, **9** (2), 1–31.

Lachowska, M., Mas, A. and Woodbury, S. A. (2018). *Sources of displaced workers' long-term earnings losses*. Tech. rep., National Bureau of Economic Research.

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

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.

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. (2020). The Costs of Job Displacement over the Business Cycle and Its Sources: Evidence from Germany.

Schwerdt, G., Ichino, A., Ruf, O., Winter-Ebmer, R. and Zweimüller, J. (2010). Does the color of the collar matter? employment and earnings after plant closure. *Economics Letters*, **108** (2), 137–140.

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.

Song, J., Price, D. J., Guvenen, F., Bloom, N. and Von Wachter, T. (2018). Firming up inequality. *The Quarterly Journal of Economics*, **134** (1), 1–50.

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

group	age	logwmin1	logwmin2	jobTenure	prodsector	fsize	propensity	n
Year = 1989								
1	37.72	4.63	4.62	1586.22	0.62	217.95	50.15	1300
2	37.76	4.63	4.62	1586.24	0.60	210.33	50.15	1300
3	37.91	4.61	4.59	1643.67	0.48	1566.20	50.08	383017
Year = 1990								
1	37.36	4.57	4.56	1771.02	0.56	116.87	50.20	1410
2	37.33	4.57	4.56	1806.26	0.57	114.40	50.20	1410
3	37.81	4.63	4.60	1870.38	0.49	1541.54	50.09	374545
Year = 1991								
1	37.00	4.57	4.54	1970.06	0.54	156.46	50.15	1236
2	36.60	4.56	4.53	1937.70	0.55	159.43	50.15	1236
3	37.70	4.65	4.62	2027.62	0.50	1672.56	50.08	383144
Year = 1992								
1	37.60	4.64	4.60	2072.85	0.42	302.98	50.24	2875
2	37.32	4.65	4.61	2087.12	0.48	265.11	50.24	2875
3	37.50	4.67	4.63	2167.48	0.51	1640.96	50.18	387729
Year = 1993								
1	37.60	4.67	4.66	2235.55	0.70	649.40	50.32	4231
2	37.88	4.66	4.65	2177.72	0.68	582.87	50.32	4231
3	37.34	4.67	4.65	2290.31	0.51	1628.39	50.26	395203
Year = 1994								
1	37.28	4.62	4.61	2097.54	0.53	75.99	50.29	1561
2	37.24	4.63	4.62	2122.31	0.53	75.75	50.29	1561

(continued)

group	*age*	*logwmin1*	*logwmin2*	*jobTenure*	*prodsector*	*fsize*	*propensity*	*n*
3	37.17	4.67	4.65	2377.48	0.51	1577.82	50.09	405626
Year = 1995								
1	37.74	4.64	4.61	2266.26	0.50	244.13	50.27	3134
2	37.50	4.64	4.61	2247.47	0.51	241.16	50.27	3134
3	37.23	4.69	4.66	2492.18	0.50	1552.11	50.19	406903
Year = 1996								
1	37.78	4.64	4.62	2355.85	0.70	192.62	50.39	3390
2	37.67	4.64	4.62	2362.40	0.73	198.06	50.39	3390
3	37.47	4.71	4.68	2631.92	0.50	1526.64	50.20	404581
Year = 1997								
1	37.76	4.64	4.63	2407.20	0.61	90.21	50.21	1353
2	37.51	4.63	4.62	2486.15	0.60	87.93	50.21	1353
3	37.52	4.71	4.70	2713.32	0.50	1462.91	50.08	410672
Year = 1998								
1	37.53	4.69	4.67	2345.60	0.45	194.64	50.16	1689
2	37.65	4.67	4.64	2390.67	0.44	186.12	50.16	1689
3	37.54	4.71	4.69	2807.54	0.50	1452.41	50.10	412663
Year = 1999								
1	37.70	4.66	4.64	2709.91	0.56	355.54	50.24	3149
2	37.56	4.65	4.63	2673.30	0.57	367.26	50.24	3149
3	37.62	4.73	4.70	2887.17	0.50	1409.91	50.19	408815
Year = 2000								
1	37.84	4.70	4.68	2382.79	0.66	156.46	50.23	1959
2	37.82	4.70	4.68	2362.34	0.66	158.08	50.23	1959
3	37.73	4.74	4.72	2947.87	0.49	1412.60	50.12	409693
Year = 2001								
1	37.43	4.63	4.63	2439.30	0.60	116.62	50.33	2389

(continued)

group	*age*	*logwmin1*	*logwmin2*	*jobTenure*	*prodsector*	*fsize*	*propensity*	*n*
2	37.44	4.63	4.62	2431.16	0.62	119.04	50.33	2389
3	37.89	4.74	4.73	3011.00	0.49	1393.34	50.14	409842
Year = 2002								
1	37.88	4.72	4.70	3000.88	0.68	573.29	50.23	3241
2	37.81	4.74	4.72	3097.73	0.70	558.25	50.23	3241
3	38.02	4.75	4.72	3042.32	0.49	1315.66	50.19	411583
Year = 2003								
1	38.85	4.60	4.59	3367.89	0.25	10152.64	54.40	6237
2	38.75	4.63	4.62	3078.66	0.22	10054.63	54.00	6237
3	38.07	4.75	4.73	3054.22	0.49	1111.77	50.25	409574
Year = 2004								
1	37.13	4.70	4.68	2479.83	0.30	193.34	50.18	1792
2	37.33	4.71	4.68	2542.28	0.34	172.37	50.18	1792
3	38.15	4.76	4.74	3082.23	0.50	1041.74	50.11	402703
Year = 2005								
1	38.01	4.62	4.63	2608.31	0.62	88.96	50.26	1391
2	38.01	4.63	4.64	2550.76	0.63	88.08	50.26	1391
3	38.21	4.76	4.75	3063.74	0.49	1083.65	50.08	407225
Year = 2006								
1	39.28	4.85	4.82	3892.02	0.30	769.82	50.20	2130
2	39.67	4.84	4.82	4111.85	0.33	797.69	50.20	2130
3	38.28	4.76	4.74	2994.88	0.47	1199.27	50.12	425399
Year = 2007								
1	37.40	4.61	4.59	2335.67	0.59	163.39	50.14	1054
2	37.84	4.60	4.59	2427.92	0.59	166.56	50.13	1054
3	38.38	4.78	4.75	2983.83	0.47	1230.65	50.06	439162

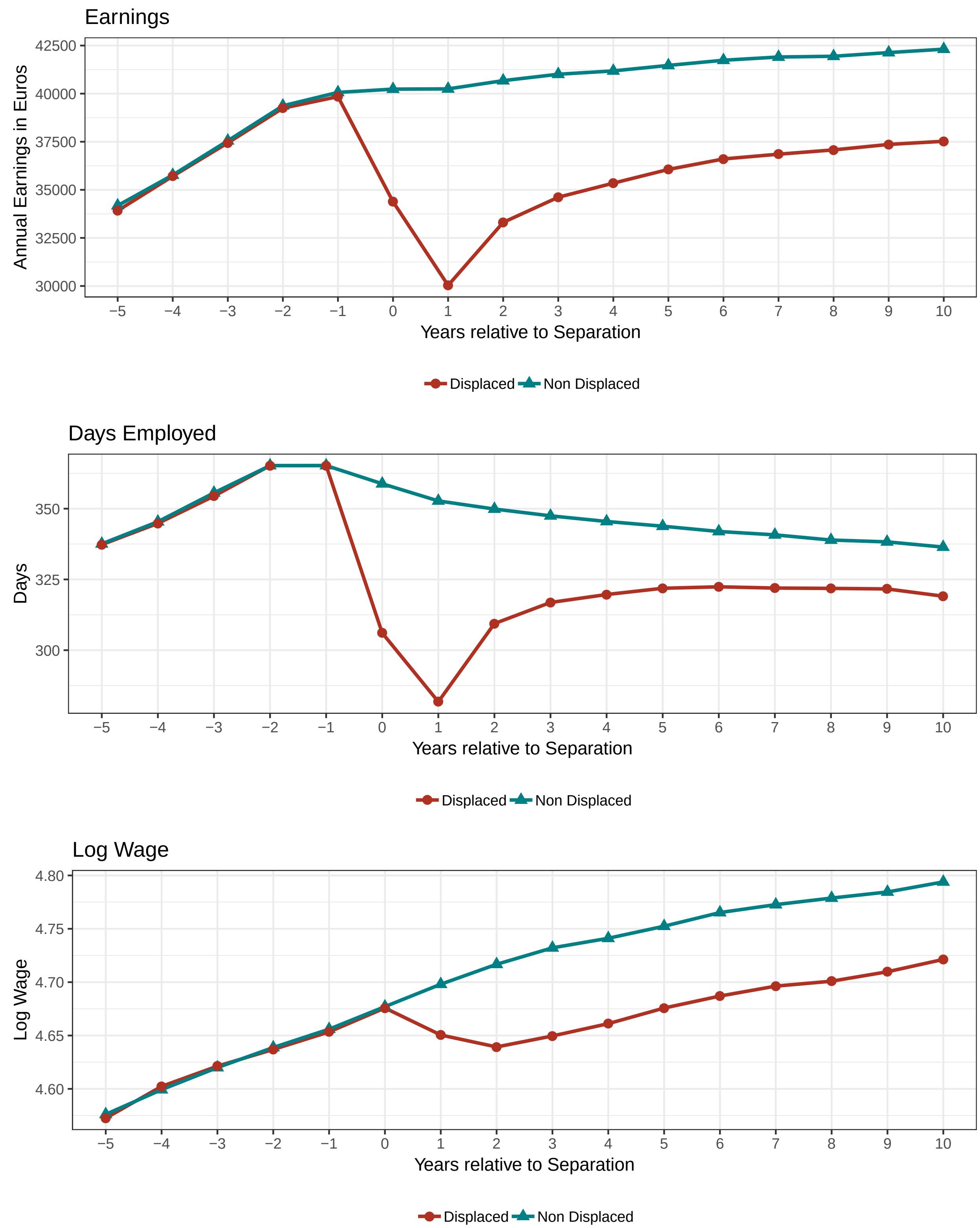


Figure 16: Earnings evolution and days employed over time for workers displaced and control group. Control group was select via propensity score matching

B. 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.[40]

A. Computation of AKM model

We follow closely Card *et al.* (2013) for the computation of the AKM model. We use the universe of male employment spells covered by social security from 1984-2017. 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 55,079,910 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 52,034,878 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 10 presents the log-wage variance decomposition based on the AKM regression results.

B. Partitioning Variables

Table 11 presents the cut points used for the categorization of continuous variables and Figure 17 shows how these variables are correlated with each other in the estimation sample.

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

Table 10: AKM Decomposition

Person and establishment parameters		
Number of person effects	3739430	
Number of establishment effects	591512	
Variance Decomposition	Total	Share
Var(Person effects)	0.077	42.5
Var(Establ effects)	0.055	30.1
Var(Xb)	0.028	15.2
Var(Residual)	0.031	17.0
2Cov(Person/establ. effects)	0.019	10.2
2Cov(perons/Xb)	-0.021	-11.5
2Cov(establ/Xb)	-0.006	-3.5
Var(log wages)	0.182	100
Summary of Estimation		
Adj R^2	0.8145	
Sample size	52,034,878	

Table 11

	P10	P20	P30	P40	P50	P60	P70	P80	P90
Firm FE	-0.21	-0.11	-.043	0.01	0.05	0.09	0.13	0.18	0.23
Firm Size	5	13	26	50	96	178	328	658	1809
Income t-1	8,199	17,603	24,011	28,322	32,164	36,277	41,382	48,578	57,514
Job Tenure	61	216	366	647	997	1461	2126	3027	4740
Herfindahl Index	0.004	0.008	0.012	0.016	0.021	0.028	0.037	0.058	0.121
Avg. Firm FE	-0.04	-0.02	0.00	0.02	0.02	0.02	0.03	0.05	0.06
Industry U-Rate	0.028	0.050	0.071	0.082	0.090	0.097	0.105	0.122	0.180
Change Ind. U-Rate	-0.108	-0.056	-0.020	0	0.001	0.026	0.059	0.104	0.228
Regional U-Rate	0.056	0.065	0.075	0.083	0.089	0.099	0.109	0.119	0.129
Change Regio. U-Rate	-0.081	-0.050	-0.027	0	0.021	0.041	0.063	0.103	0.148
Firm Sep. Rate	0.081	0.111	0.136	0.162	0.190	0.226	0.278	0.366	0.569

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

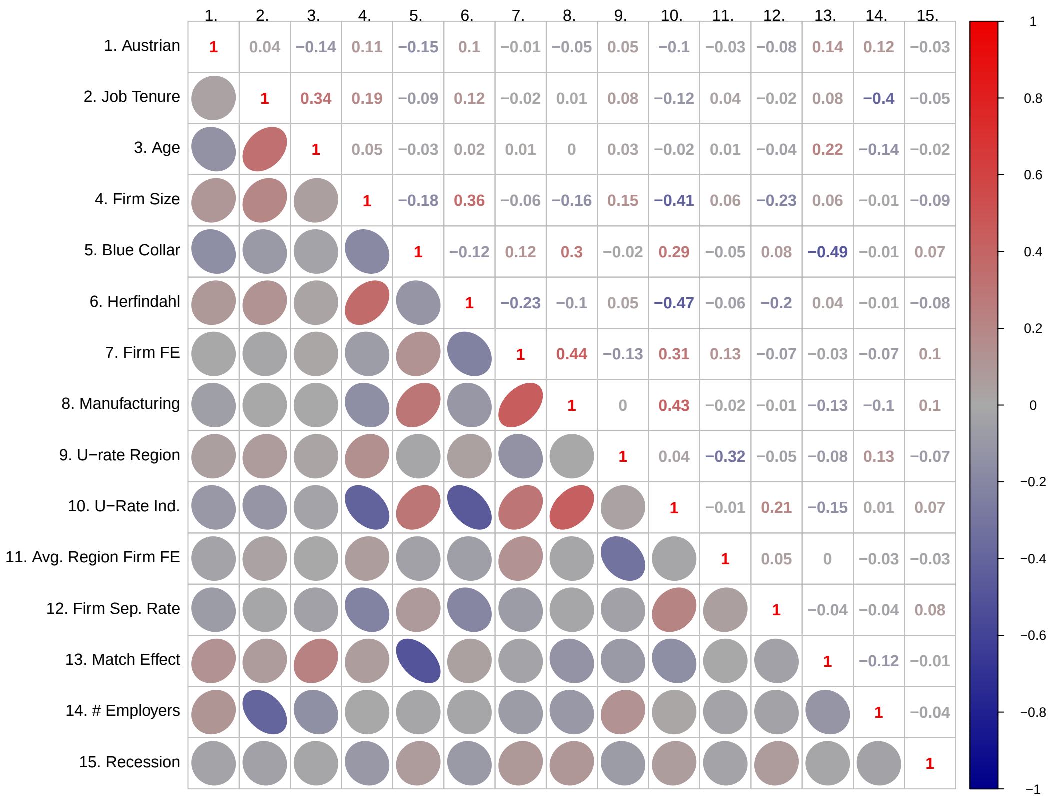


Figure 17: Correlogram of partitioning variables

C. Marginal Effect and Partial Dependence Plots

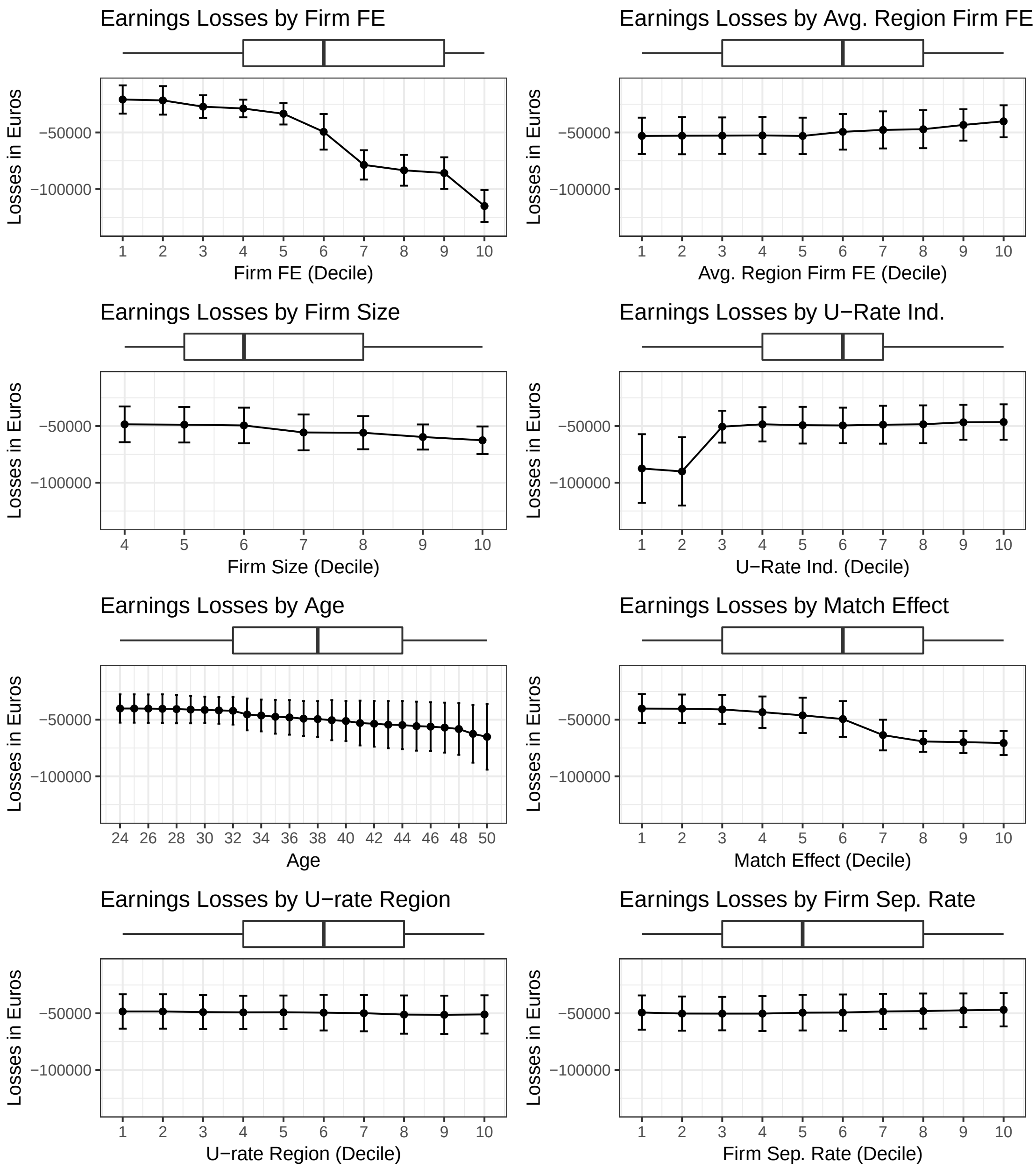


Figure 18: GRF estimates of the cumulative 11-year earnings losses with 95% confidence intervals (part I). All variables that are not used in a given plot are set to their mean values. The y-axis indicates cumulative 11-year earnings losses, while the x-axes depict global deciles of the unrestricted dataset (except for the age, number of employees, blue collar, manufacturing, and Austrian where the actual values are used). The boxplots above figures represent the distribution of the partitioning variable in the restricted dataset (for the details see Subsection B of Section II).

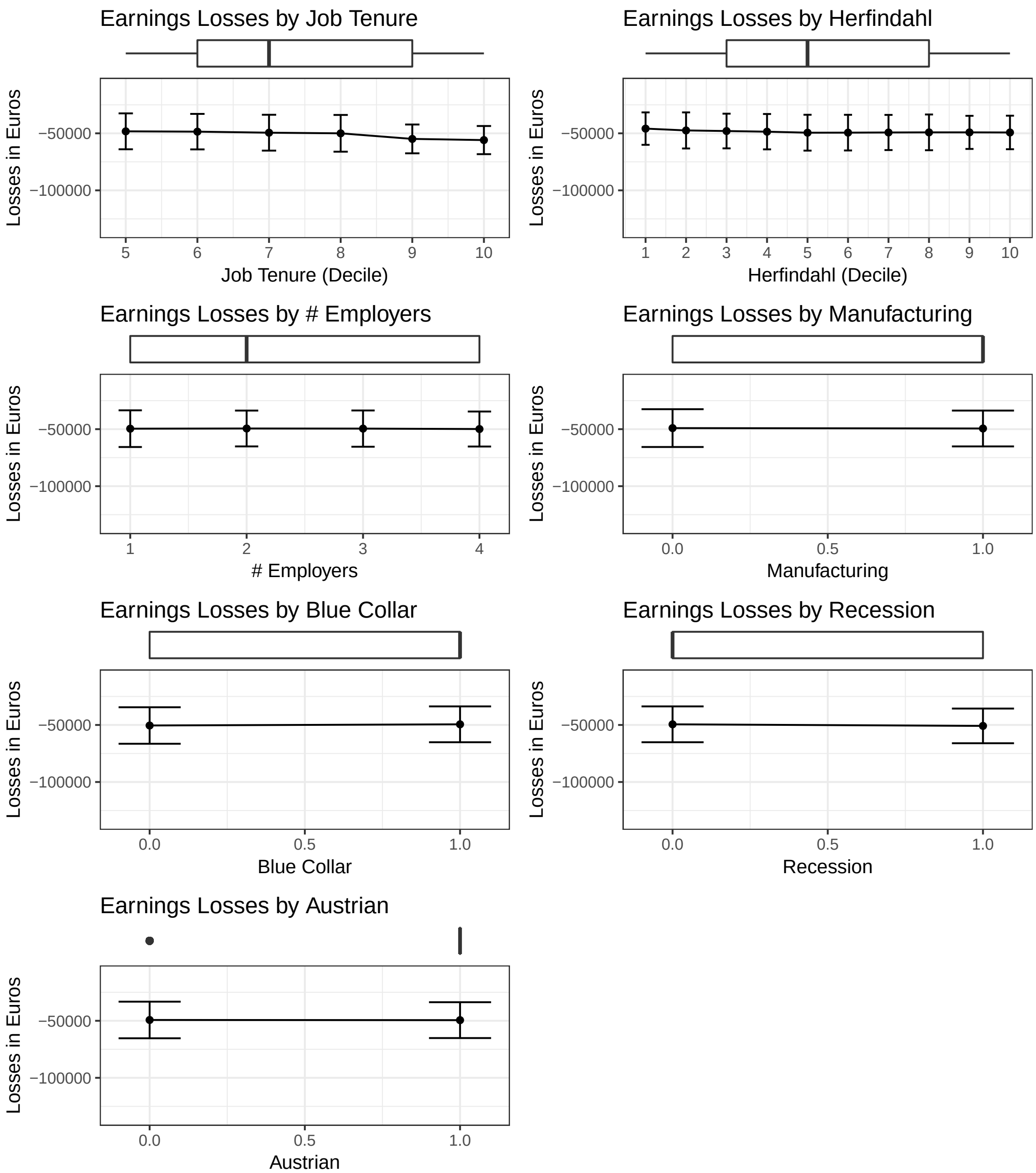


Figure 19: GRF estimates of the cumulative 11-year earnings losses with 95% confidence intervals (part II). All variables that are not used in a given plot are set to their mean values. The y-axis indicates cumulative 11-year earnings losses, while the x-axes depict global deciles of the unrestricted dataset (except for the age, number of employees, blue collar, manufacturing, and Austrian, where the actual values are used). The boxplots above figures represent the distribution of the partitioning variable in the restricted dataset (for the details see Subsection B of Section II).

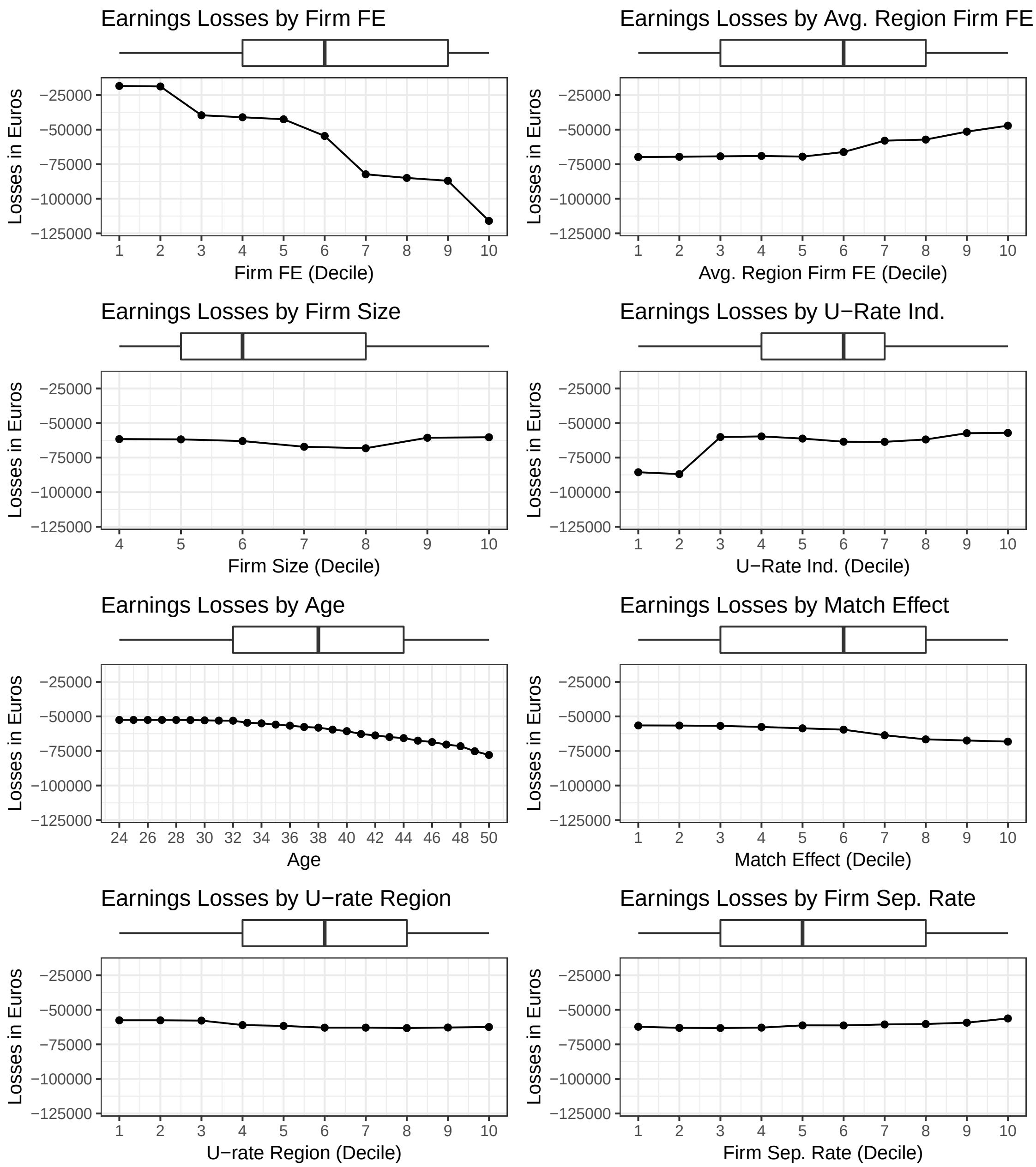


Figure 20: The figure shows partial dependence plots (for details on computations see Subsection C of Section VI) for different variables (part I). The y-axis indicates cumulative 11-year earnings losses, while the x-axes depict global deciles of the unrestricted dataset (except for the age, where the actual values are used). The boxplots above figures represent the distribution of the partitioning variable in the restricted dataset (for the details see B).

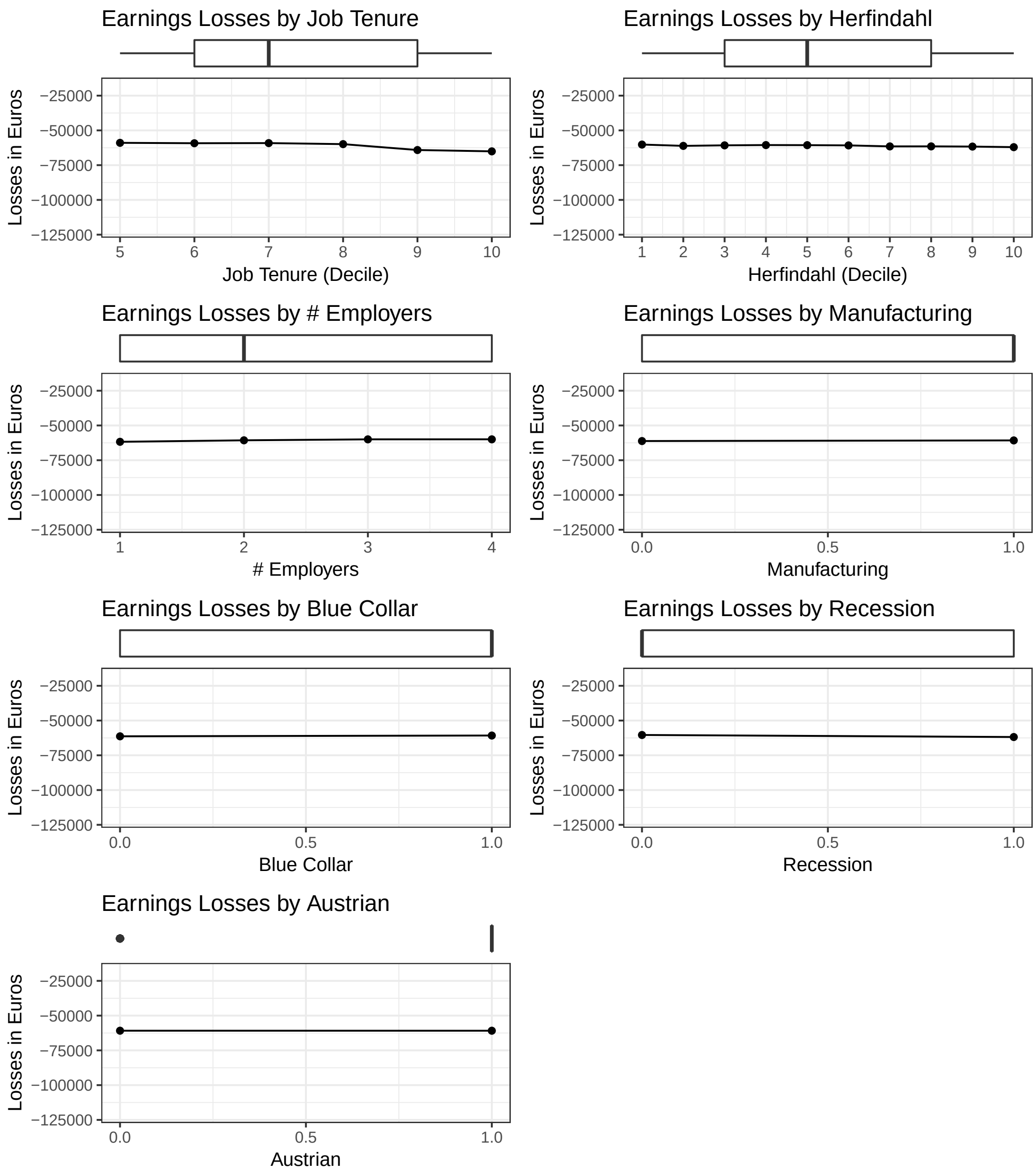


Figure 21: The figure shows partial dependence plots (for details on computations see Subsection C of Section VI) for different variables (part II). The y-axis indicates cumulative 11-year earnings losses, while the x-axes depict global deciles of the unrestricted dataset (except for the age, where the actual values are used). The boxplots above figures represent the distribution of the partitioning variable in the restricted dataset (for the details see B).

D. The sources of log wage losses

In this section we present results for the random forest grown to maximize the heterogeneity in changes in log wages. Those changes are a measure of percentage losses in terms of pre-displacement wages.

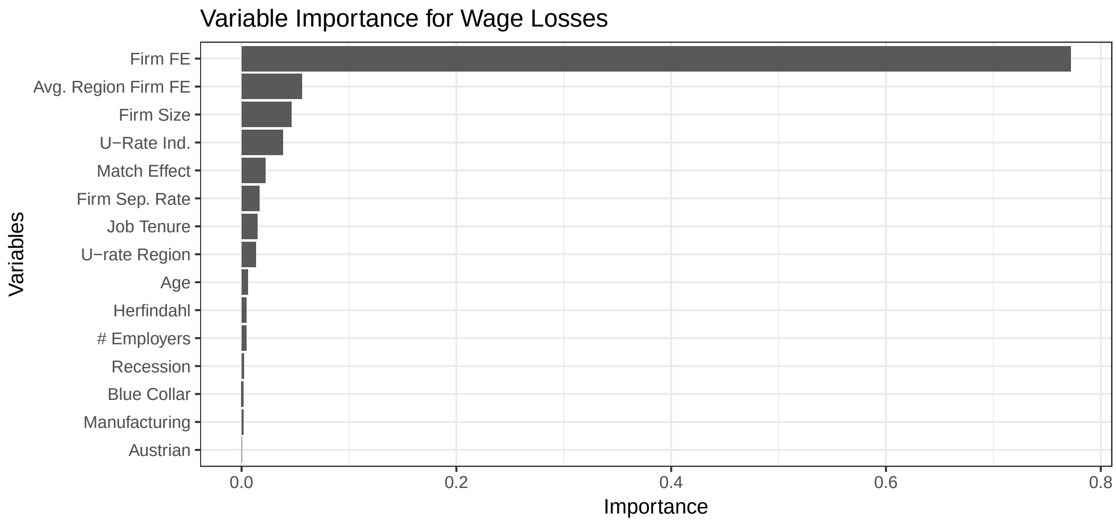


Figure 22: Variable importance for the log wage losses. Depth-adjusted variable frequency in splits in the GRF with a decay exponent equal to -2 and the maximum depth level of nodes equal to 4. All values sum to 1.

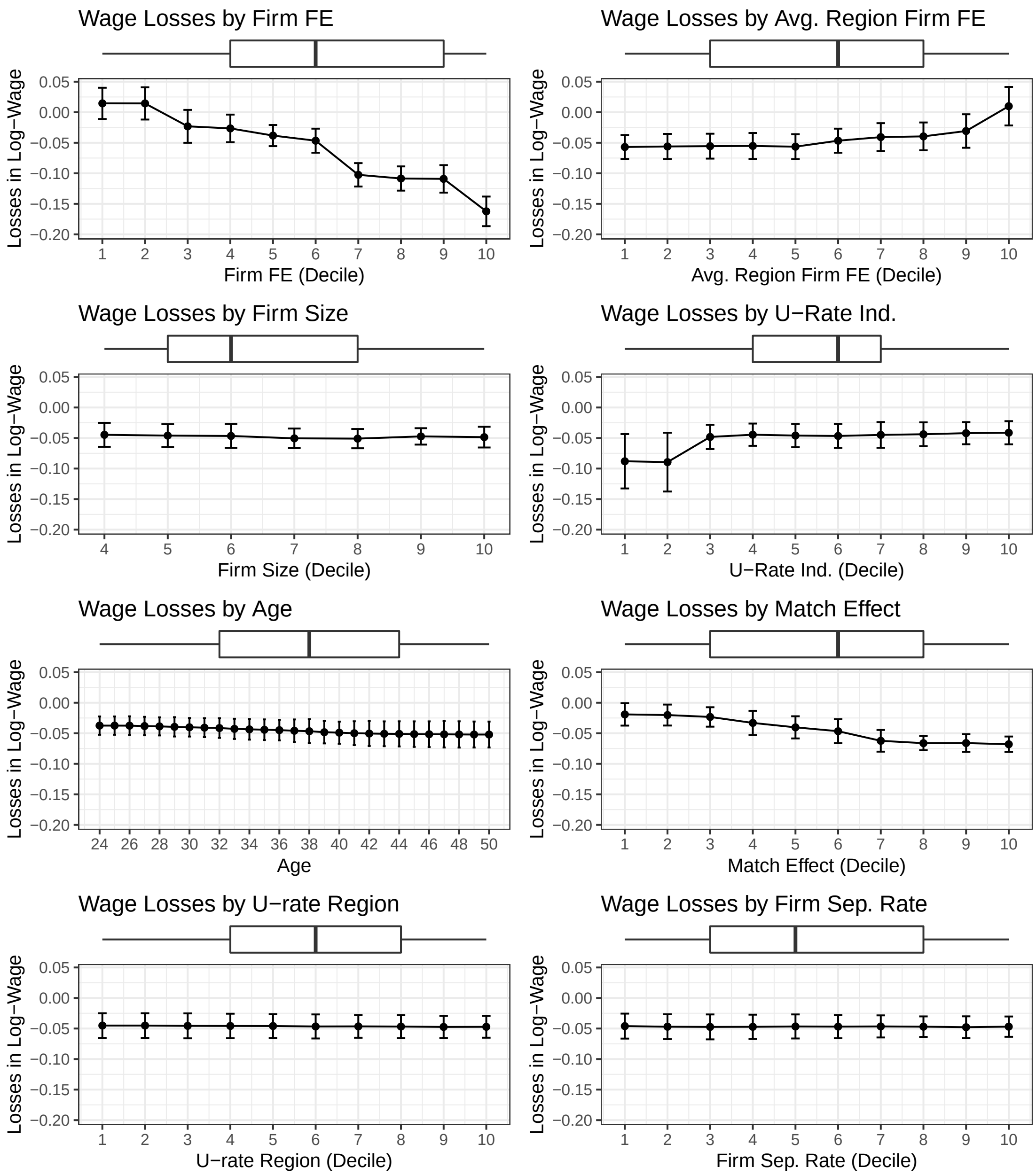


Figure 23: GRF estimates of the log wage losses with 95% confidence intervals (part I). All variables that are not used in a given plot are set to their mean values. The y-axis indicates cumulative 11-year earnings losses, while the x-axes depict global deciles of the unrestricted dataset (except for the age, number of employees, blue collar, manufacturing, and Austrian where the actual values are used). The boxplots above figures represent the distribution of the partitioning variable in the restricted dataset (for the details see Subsection B of Section II).

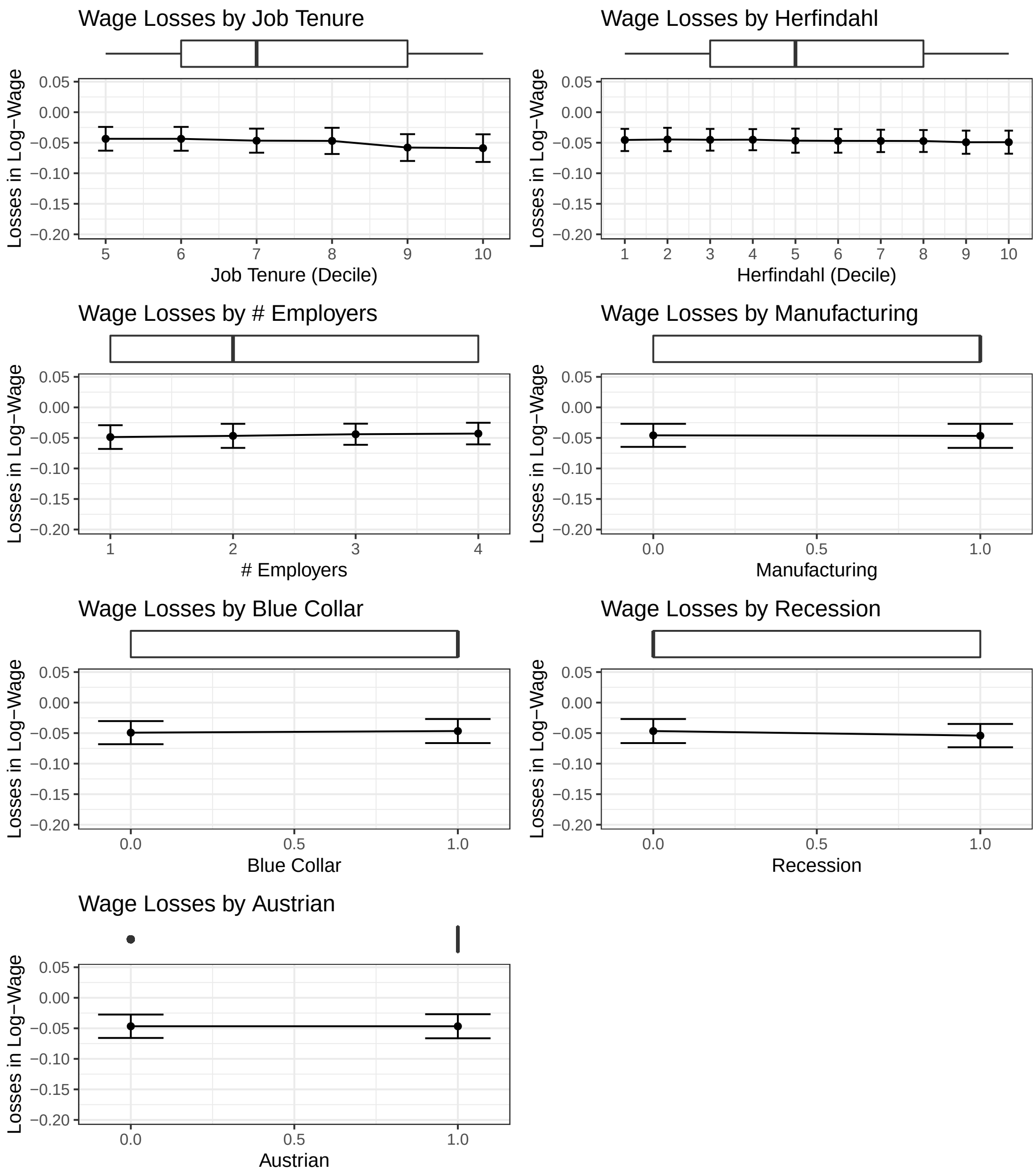


Figure 24: GRF estimates of the log wage losses with 95% confidence intervals (part II). All variables that are not used in a given plot are set to their mean values. The y-axis indicates cumulative 11-year earnings losses, while the x-axes depict global deciles of the unrestricted dataset (except for the age, number of employees, blue collar, manufacturing, and Austrian, where the actual values are used). The boxplots above figures represent the distribution of the partitioning variable in the restricted dataset (for the details see Subsection B of Section II).

E. Robustness: Firm Fixed Effect

We perform the following robustness exercise. In the baseline specification, we computed the firm fixed effect measure using the universe of all male observations except the observations from our earnings loss sample. This means that however the masslayoff affected labor market outcomes of the individuals in our sample, it cannot have an effect on the measurement of the firm fixed effect, as these individuals were excluded in the AKM regression. But Post-masslayoff of those workers that stayed at mass layoff firms, but were not selected by the propensity score matching could influence the estimation of firm fixed effects and lead to a mis-measurement of the firm pay premia at the time of separation.

To rule out that any of our results are driven by this, we re-estimate the AKM regression equation (4) only on observations before 1998. We then re-estimate our forest with workers after 1998. Figure 25 shows that the estimated importance of firm fixed effects is essentially unchanged compared to our baseline specification. Furthermore we recompute the marginal effects for all partitioning variables. Figures 26 and 27 show that these are also very similar to our baseline results. Some changes are to be expected, as we also consider only a subperiod. The main result is essentially unchanged, workers from low fixed effect firms still face considerably lower earnings losses than workers employed at high fixed effect firms. It is still the variable that shows the highest slope in earnings losses.

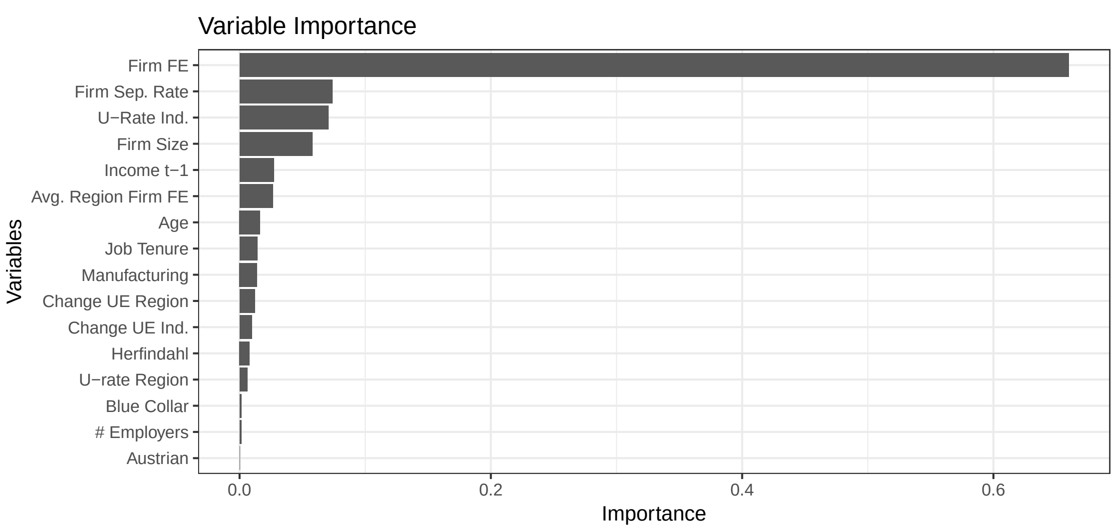


Figure 25: Variable importance using only workers from 1999 or later, while firm fixed effect is computed on observations before 1998. Depth-adjusted variable frequency in splits in the GRF with a decay exponent equal to −2 and the maximum depth level of nodes equal to 4. All values sum to 1.

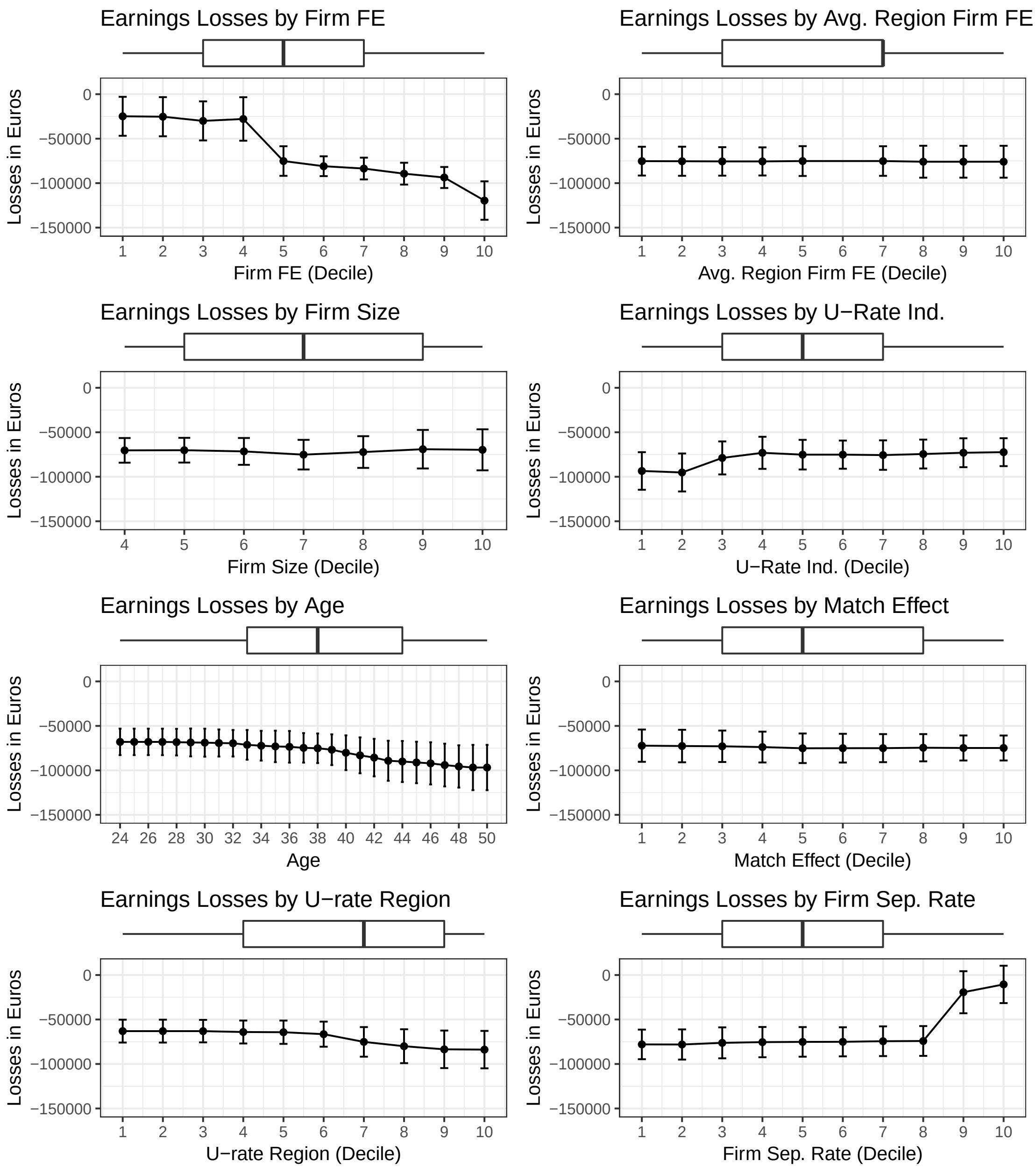


Figure 26: GRF estimates of the cumulative 11-year earnings losses with 95% confidence intervals (part I), only workers from 1999 onwards. Firm fixed effects are computed on observations before 1998. All variables that are not used in a given plot are set to their mean values. The y-axis indicates cumulative 11-year earnings losses, while the x-axes depict global deciles of the unrestricted dataset (except for the age, number of employees, blue collar, manufacturing, and Austrian where the actual values are used). The boxplots above figures represent the distribution of the partitioning variable in the restricted dataset (for the details see Subsection B of Section II).

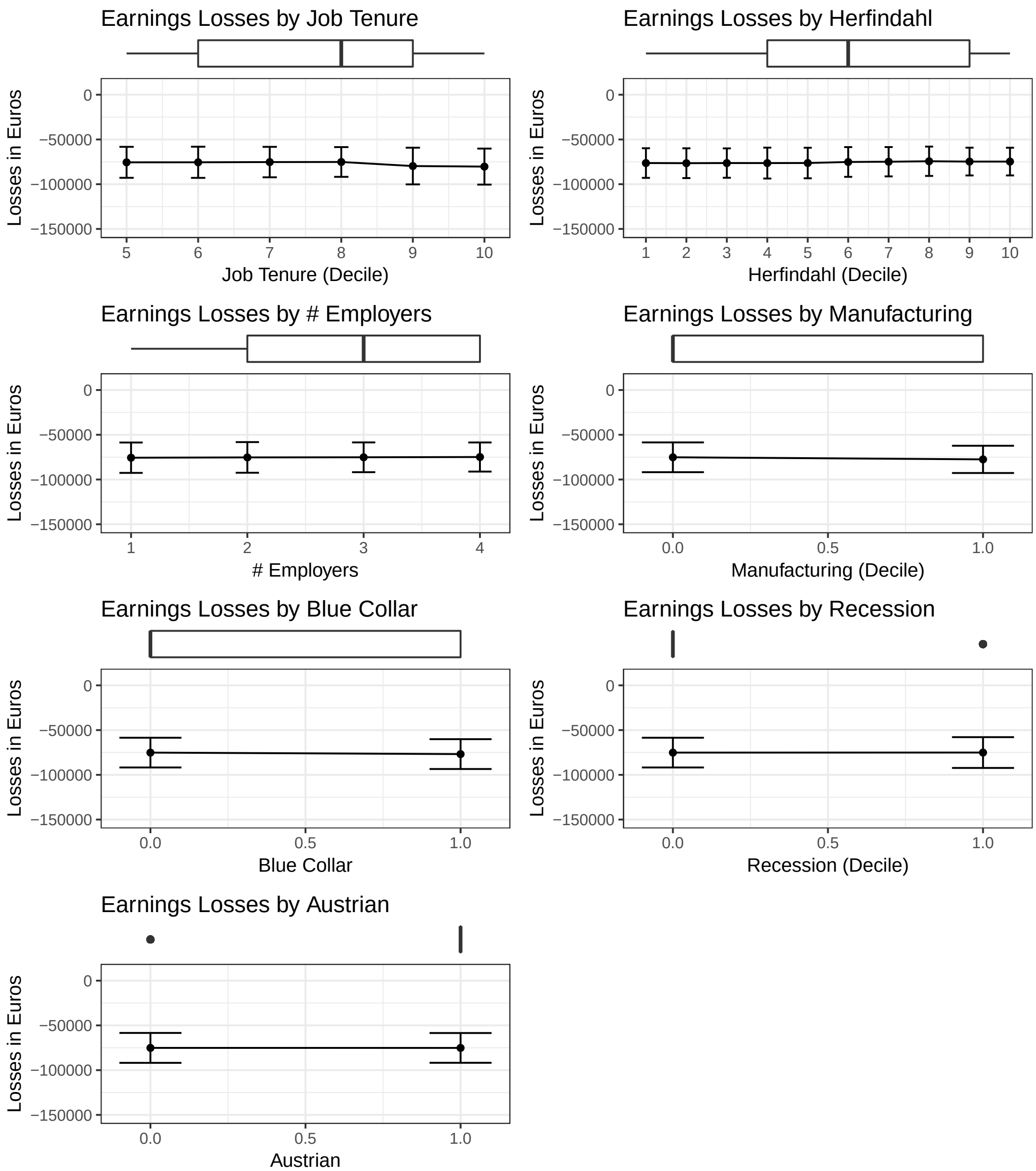


Figure 27: GRF estimates of the cumulative 11-year earnings losses with 95% confidence intervals (part II), only workers from 1999 onwards. Firm fixed effects are computed on observations before 1998. All variables that are not used in a given plot are set to their mean values. The y-axis indicates cumulative 11-year earnings losses, while the x-axes depict global deciles of the unrestricted dataset (except for the age, number of employees, blue collar, manufacturing, and Austrian, where the actual values are used). The boxplots above figures represent the distribution of the partitioning variable in the restricted dataset (for the details see Subsection B of Section II).