UNIVERSITÄT
BERN

Faculty of Business, Economics
and Social Sciences

**Department of Economics**

# Is the Impact of Opening the Borders Heterogeneous?

Costanza Naguib

23-12

December, 2023

# DISCUSSION PAPERS

**Schanzeneckstrasse 1**
**CH-3012 Bern, Switzerland**
***http://www.vwi.unibe.ch***

# Is the Impact of Opening the Borders Heterogeneous?

Costanza Naguib*

**Abstract**

We analyze the impact of an inflow of foreign workers on the wage distribution of residents in a small open economy like Switzerland. We exploit the fact that Swiss mobility regions were differently affected by the intensity and the timing of the Agreement on the Free Movement of Persons, depending on their distance from the national border. We extend the results by Beerli, Ruffner, Siegenthaler and Peri (2021) by analyzing heterogeneity in treatment effect via causal forests. We find statistically significant evidence that the treatment effect of opening the borders has been heterogeneous across age, education, and type of activity groups.

Keywords: wage distribution, Bilateral Agreements, causal forest, Conditional Average Treatment Effect
JEL codes: C14, J31

# Introduction

[1] The aim of this paper is to analyze the heterogeneous impact of a progressive opening of the labor market to an inflow of foreign workers on residents' wages[2]. The impact of migration on labor market outcomes is likely to be heterogeneous (see e.g. Dustmann et al. (2012), Dustmann et al. (2017), Llull (2018), and Beerli, Ruffner, Siegenthaler and Peri (2021, BRSP

*University of Bern, Department of Economics (VWI), Bern, Switzerland. costanza.naguib@vwi.unibe.ch

[1] I would like to thank Mirjam Bachli, Andreas Beerli, Giacomo De Giorgi, Michael Lechner, Frederic Robert-Nicoud, Michael Siegenthaler, Anthony Strittmatter, Teodora Tsankova, the participants to the SEW research seminar on 4-6 June 2018, to the SEW research seminar on 22 October 2018, to the research seminar at the University of Geneva on 8th February 2019, to the YSEM conference in Zurich on 14-15 February 2019, to the brown bag seminar in St. Gallen on 20th Febraury 2019, to the SSES conference in Geneva on 13-14th June 2019, and to the research seminar at KOF on 27th November 2019 for their insightful advice.

[2] Residents include both natives and foreigners who had previously immigrated.

henceforth)). The causal forest algorithm developed by Wager and Athey (2018) allows to estimate the Conditional Average Treatment Effects (CATEs) for each individual in the sample. Applying this machine learning algorithm, we are able to replicate and to extend BRSP's results. Estimators of treatment effect based on standard difference-in-difference designs have been shown to be potentially invalid in presence of treatment effect heterogeneity across groups (De Chaisemartin and d'Haultfoeuille (2020)). In the empirical part, we document that treatment effect is indeed heterogenous in our framework.

The machine learning algorithm also yields a measure of variable importance, i.e. it tells us which covariates explain the largest shares of the variance in the outcome variable. Therefore, we study heterogeneity in treatment effect across age, education, size of the firm and type of activity groups, as they appear to be the main drivers of wages' variance. The analysis of the differential impact of a migration reform can provide the basis for a targeted policy response to the (potentially) adverse effects of the reform itself. We contribute to the small but growing literature which applies machine learning methods to economic research questions (see e.g. Bertrand et al. (2017), Davis and Heller (2017) and Kleinberg et al. (2017)).

In the wake of Gulyas and Pytka (2020), we apply the generalized random forest by Athey et al. (2019) to a difference-in-difference setup, in order to estimate the heterogeneous impact of a migration flow on wages, conditional on a large number of observables. We build on existing research by exploiting information from a quasi-natural experiment, i.e. the entry into force of the Agreement on the Free Movement of Persons between Switzerland and the European Union, which took place on 1st June 2002. This reform liberalized the status of both foreign migrant workers and cross-border workers, where the latter gained geographical mobility within the Swiss Border region. Cross-border workers represent a particularly interesting case, since they are going to influence labor market conditions by an increase in labor supply. However, they are not going to influence, say, house prices or rent, since they are not migrating into the country in which they work (BRSP 2021). This allows us to assume that general equilibrium effects on markets other than the labour market (e.g. house prices or local consumption) have been negligible (BRSP 2021).

The present paper is one of the few studies exploiting changes in policies for cross-border workers in order to study the impact of migration on labor market outcomes. Dustmann et al. (2017) analyze the strong increase in employment of Czech commuters in regions close to the border shortly after the opening of Germany's labor market to Czech workers in 1991. However, in that case, the policy was unannounced and temporary, and it was reversed in 1993. On the contrary, the Swiss change in the migration policy was announced in 1999 and constituted a gradual and permanent shift in migration rules for workers. As explained in Section 1.2, the opening of the borders was step-wise. Therefore, the reform is conceptually different from a large one-off labor supply shock (unlike, for instance, the Mariel boat-lift, where a large number of new immigrants were suddenly placed in Miami's labor market (Card (1990)). In a work which is strictly connected to ours, Beerli, Ruffner, Peri and Siegenthaler (BRSP 2021) focus on mean impact estimations of the same Swiss migration reform on wages, employment and innovation. We extend their results on wages by allowing for heterogeneity in treatment effect. As far as the external validity of the estimated causal effects is concerned, according to BRSP 2021, this change in policy was permanent (and perceived as such), in contrast with other cases of changes in the migration policy which were only temporary.

Only a few papers have tackled heterogeneity in treatment effect by now. Relevant examples in this literature are Akgündüz and Torun (2020), who find that the impact of a large inflow of Syrian refugees in Turkey has been heterogenous across a variety of dimensions, among which age, educational level and firm size. Further, Mitaritonna et al. (2017) study the impact of an increase in the local supply of immigrants on firms' outcomes, allowing for heterogeneous effects across firms in the case of France. Dustmann et al. (2012) studies the impact of immigration across the distribution of wages and find evidence of a negative impact in the bottom 20% only.

From a theoretical viewpoint, we expect to find heterogeneity in treatment effects along the following dimensions. First, young workers should be more able to adapt to the new working conditions and to learn new skills if necessary to face competitions from the immigrants, more than their older colleagues. Hence, we expect to find a positive or zero treatment effect for younger worker and we could potentially find a negative impact for older workers. Further, we expect more skilled and highly educated workers to be less sostituable with foreigners and hence to be strategic complements, rather than strategic substitutes with foreign workers. In particular, we expect to find a positive treatment effect for highly educated individuals (e.g. with a university or a professional university degree) and a negative of zero effect for those with lower educational qualifications (e.g. compulsory education only). These results would be consistent with the findings by Beerli, Peri, Ruffner and Siegenthaler (2021). For a similar reason, we expect workers who execute tasks with a high degree of autonomy and/or with management responsibility to be positively impacted by the free movement reform, whereas those who only execute simple and repetitive tasks and do not have any management responsibility may be negatively impacted by it, since now they have to face an increased competition (see e.g. Dustmann et al. (2012), Beerli, Peri, Ruffner, Siegenthaler (2021)). Further, we deem that larger firms are more likely to be affected by inflows of foreign workers. We expect individuals with more years of service in the same firm to have a positive (or zero) treatment effect, whereas those with less experience may witness a negative treatment effect, for reasons similar to those mentioned above for the variables education and skill requirements of the job. Finally, we expect to find heterogeneity in treatment effect along the dimension of the type of activity of the worker. Just to make an example, sectors like education, being mostly country-specific, are likely to have record a zero treatment effect. Other industries, such as research and development, in which natives and foreigners are expected to be strategic complements, should exhibit a positive treatment effect. On the other hand, in industries like construction or manufacturing, native and foreigners are more likely to be strategic complements, hence we expected the estimated treatment effect to be negative (or zero).

Estimating CATEs may reconcile existing puzzles, i.e. why some studies found a negative impact and some other a positive impact of foreign inflows on natives' labor market outcomes (see e.g. Borjas (2003), Borjas et al. (2008), Card (2009)). The mean impact has been proven to mask important heterogeneity in treatment effects (see e.g. Bitler et al. (2006)). Simply subsetting the original dataset (e.g. into those with a university degree only, those aged 25-35 only, and so on) and then applying standard difference-in-difference estimation to each of the subsamples may be an alternative. However, in this latter case the selection of which variables are deemed to be relevant drivers of heterogeneity in treatment effects, as well as of which are the relevant cutting points (e.g. for age) would be arbitrarily determined by the researcher[3]. We prefer to maintain an agnostic approach and let the data speak, by including in principle all exogenous explanatory variables which are available in the data and then letting the algorithm select which of them are the main drivers of heterogeneity in treatment effect. Once we know this, we are able to estimate and compare the CATEs among the dimensions across which the most differ.

We find evidence of statistically significant heterogeneity in treatment effect across workers' characteristics. The main drivers of heterogeneity in treatment effect are age, education, size of the firm and type of activity (industry). We find a positive impact of the reform for young workers (i.e. in their 20s and 30s), as well as for those working for a large firm. However, the estimated CATEs are notably small, as well as their variations across individuals. Differ-

[3]Standard difference-in-difference results are consistent with our machine learning results. These results are not included here for reason of space, but are available from the author upon request.

ences in CATEs among workers with different characteristics are around 0.5-1 Swiss franc per month. Hence, we deduce that, while heterogeneity in treatment effect is statistically significant in this framework, it is nevertheless not relevant for policy-making decisions. BRSP 2021 also analyze the impact of the reform on total local employment and they find no statistically significant impact. This is true both in their full sample and in their subsample of highly educated workers. For this reason, we refrain from studying the heterogeneous impact of the reform on employment. Further, the data that both us and BRSP 2021 use only include information about employed individuals. Hence, an analysis of unemployment is only possible at the aggregate (local) level, as BRSP 2021 did. However, an analysis of the heterogeneous determinants of the individual probability of being employed is not feasible with the available data. In the present paper we define natives as Swiss nationals and foreigners as those with nationality other than Swiss. Individuals with double nationality (both Swiss and foreigner) are classified as natives.
The remainder of this paper is organized as follows. Section 1 presents an overview of the related literature and of the institutional framework. Section 2 describes the method used for estimation, whereas Section 3 is devoted to the description of the dataset. In Section 4 the estimation results are presented and discussed. Section 5 concludes. Appendix A provides further details on the institutional reform at hand and Appendix B presents additional data analysis.

# 1 Literature review

## 1.1 Relation to existing literature

There is a broad literature on the influence of immigration on wages and employment. Most of the studies focus on the mean estimation of the impact, thus neglecting the potential dramatic heterogeneity of the treatment effect of being exposed to a sudden increase in migration flows of foreigner workers. Borjas (2003) and Borjas et al. (2008), as well as Bratsberg and Raaum (2012) find a negative impact of migration on natives' wages. On the contrary, Card (2009), Peri and Sparber (2009), Peri and Sparber (2011), Ottaviano and Peri (2012), Peri et al. (2015), Foged and Peri (2016) find, if any, a positive impact of foreign workers inflows on natives' wages, thus suggesting that foreign and native workers are strategic complements rather than strategic substitutes. Further, contrary to the common belief that immigrant inflows cause outflows of native workers, Card and DiNardo (2000) and Peri (2012) find evidence that increases in immigrant population in certain skill groups leads to a slight increase in the native population in the same skill-group, i.e. there is no crowding-out. An increasing number of studies exploits a quasi-experimental framework to assess the impact of immigration on wages (see e.g. Card (1990), Friedberg (2001), Kugler and Yuksel (2008), Glitz (2012), Prantl and Spitz-Oener (2019)). In all these cases the analysis relies on an unexpected change in the migration regulations, which could not be anticipated by the labor market agents. In our context, on the contrary, the change in the migration policy was announced and gradually introduced over the years 1999-2007, as it will be explained in Section 1.2. As far as the Swiss case is concerned, Favre (2011), Basten and Siegenthaler, and BRSP 2021 find that the impact of the migration reform on wages has been, if any, positive, whereas there has been no statistically significant impact on aggregate employment. Native and foreign workers are strategic complements in this framework. Both Favre (2011) and BRSP 2021 attempt an analysis of heterogeneity in treatment effect by running the analysis on specific subsamples. They find hints of a positive impact on wages for high educated workers and zero impact for low educated ones. These findings motivates our research question. In the empirical part, we confirm and extend BRSP 2021 results. In a work closely connected with ours, Poulos et al. (2021) study the effect of the Free Movement of Persons and the Schengen Agreement on the share of cross-border workers in sending border regions, using European Labor Force Survey data aggregated to the region level and a matrix completion estimation approach. Their results show that opening the border increases the share of non-resident workers in host countries.

## 1.2 The institutional framework

In 1999, the first round of the bilateral agreements between Switzerland and the EU was signed. The Bilateral Agreements I included the Agreement on the Free Movement of Persons, which lifted previous restriction on the access to the Swiss labor market for foreign workers. Between 1999 and 2007, Switzerland progressively opened its borders to both foreign workers coming to live in Switzerland and to cross-border workers. Cross-border workers are residents of the bordering countries (Italy, Germany, Austria and France) that are allowed to work in Switzerland, commuting from their countries of residence. As far as cross-border workers are concerned, before 2002 only individuals living within a radius of 20 km from the Swiss border were allowed to enter the Swiss labor market as cross-border workers and their job permit was only valid for a single Canton. After 2002, people coming from every part of the neighbouring countries could qualify as cross-border workers, and they could move within the whole Swiss border region, but not outside it. In the first years after the Bilateral Agreements came into force (2002), this liberalization was mitigated by the priority requirement. This means that, before hiring a foreign cross-border worker, firms had to show to the cantonal migration office that they had not been able to find a worker with the desired characteristics within a reasonable time frame. This requirement was abandoned in 2004, and the concept of Swiss border region itself was abolished in 2007. As a consequence of the above-mentioned liberalization of the definition of cross-border workers, their number in the Swiss border region relevantly increased between 1998 and 2010, going from 7% to 9.5% of total employment (BRSP 2021). The other foreign workers were still subject to both quotas and the priority requirement until 2004. Their status was fully liberalized afterwards. The entry into force of the Agreement on the Free Movement of Persons provides an adequate framework for an impact evaluation study. This agreement fully and permanently removed previous restrictions to the Swiss labor market. Further, the timeline of the events is clear-cut: before 1999 there was no hint that Switzerland was going to pursue the policy of opening its borders to foreign labor force, between 1999 and 2002 there was an anticipation phase (i.e. before the agreement actually entered into force), between 2002 and 2004 there was a transition phase, with the above-mentioned quotas and priority requirement still being and place, and from 2004 onwards there was a full liberalization of both foreigners and cross-border workers (even if the latter ones were still restricted to work exclusively in the border region until 2007). In this framework, the Swiss mobility areas (i.e. commuting zones) were exposed to inflows of foreign workers of different intensities, depending on their distance from the national border. The border region, in particular, was contemporaneously exposed to both flows of cross-border workers and to flows of foreigners migrating in Switzerland. This provides us with a treatment group (i.e. areas very close to the border) and a control group (i.e. areas farther away from the border but similar to the former ones).

# 2 The empirical framework

## 2.1 Setup

We are interested in estimating a model of the following form:

$$y_{it} = X_{it}'\beta + \gamma_0 Year_t + \gamma_1 d_i + \gamma_2 d_i * Year_t + \varepsilon_{it} \tag{1}$$

where the outcome variable $y_{it}$ stands for log hourly real wage, $X_{it}$ is a vector including all explanatory variables available in our dataset, which is: level of qualifications required for the job[4], number of individuals employed by the firm (size of the firm), dummies for educational level, professional position (i.e. with or without management responsibilities), years of service, gender, age, type of work contract (public law contract, collective agreement, firm-level agreement or individual private law contract), type of activity of the firm (NOGA two digits). The only variable that is available in the dataset but we exclude from the covariate vector is civil status, due to its potential endogeneity/dependence on wage. Only individuals working in the private sector are considered. $Year_t$ is dummy variable taking value 1 if the date of $t$ of the observation is later than the start of the treatment period and zero otherwise. As explained in Section 1.2, the Agreement on the Free Movement of Persons entered into force on 1st June 2002, however, the full liberalization phase only started in 2004, when quotas and the priority requirement were abolished. Hence, in the following we consider 2004 as the starting date of the treatment[5]. $d_i$ is the variable representing the treatment. It takes value one if an observation is recorded to be closer than 15 minutes of travel time (by car) to the closest national border crossing and zero otherwise. This cutoff has been chosen in the wake of BRSP 2021. Indeed, the authors classify as highly treated the observations between 0 and 15 minutes of distance from the national border. Further, they find that most of the treatment effect essentially took place in within this distance bin[6].

As in the classical difference-in-difference estimation strategy, we include interaction terms between the treatment variable and the time dummy, $d_i * Year_t$. Our main coefficient of interest is $\gamma_2$, which stands for the (average) treatment effect. Finally, $\varepsilon_{it}$ is the residual term. However, our goal is to identify the heterogeneity in treatment effects. Hence, in place of the standard diff-in-diff, we adopt a machine learning procedure built on the methodology of generalized random forests developed by Athey et al. (2019).The advantages of this approach are the following: (i) first, estimation of equation (1) does not provide any hint of underlying heterogeneity of the treatment effect. This should be particularly of interest for policy makers, which could identify which groups of resident workers (if any) have recorded wage losses due to migration inflows and could predispose policies to help them. (ii) Second, heterogeneous treatment effects are typically identified through sample splitting and estimating the model in separate data bins. While this is correct for identifying the existence of heterogeneity, however, it does not tell us

[4]This variable is coded from 1 to 4, where 4 stands for repetitive tasks only and 1 stands for the maximum level of task complexity.

[5]In a robustness check, we will consider 2002 as the starting date of the treatment. Our results are robust to this alternative model specification.

[6]As a robustness check, in the empirical part we present additional results which have been obtained by setting the cutoff point at 20 minutes of distance from the national border.

what are its driving channels. On the contrary, as it will be detailed later, the machine learning approach adopted here allows to assess which explanatory variables have the highest importance in determining the heterogeneous treatment effect. Hence, we estimate a version of equation (1) with heterogeneous treatment effects, which are conditional on the explanatory variables $X_{it}$:

$$y_{it} = X_{it}'\beta + \gamma_0 Year_t + \gamma_1 d_i + \gamma_{2i}(X_{it}) d_i * Year_t + \varepsilon_{it} \quad (2)$$

From a theoretical viewpoint, we expect individuals with the lowest educational and management responsibility levels to report a negative impact of the treatment, since they can easily be substituted by foreign workers who may even have a higher educational level and be willing to perform the same task for the same wage. On the other hand, we expect resident workers with the highest levels of formal education and management responsibility to record a (slightly) positive treatment effect on their wages. This is because we expect highly educated residents and foreign workers to be complement inputs in the firm's production function. This further motivates our choice of estimating an heterogeneous treatment effect. As in every diff-in-diff estimation, there is potential concern that the observed treatment effects may be caused by a failure in the common trend assumption across regions or by unobserved confounders. Potential confounders are changes in regional policies, unobserved demand, and productivity shocks that have a regionally unequal impact, e.g., due to differences in the industrial composition of regions. In their 2021 paper, Beerli, Peri, Ruffner and Siegenthaler have already addressed this concern in several ways, for the same period and same dataset of my analysis. Note that their main outcome variable is wage, exactly as in my analysis. Beerli, Peri Rufner and Siegenthaler (2021) show that their estimated labor market effects are robust to the inclusion of a Bartik (1991) index which accounts for the different industry composition at the regional level (it controls for region-specific, sector-driven demand trends or shocks). Similarly, their results are robust to controlling for unobserved region-specific shocks using NUTS-II regions times year fixed effects, to excluding or including the linear time trend and to dropping the most important Swiss cities one by one. Given that we use the same dataset and we consider the same time frame as BPRS (2021), we do not report here such checks[7]. In the section on robustness checks, we report the results of treatment effect estimation on the placebo years (1994-1998). The results of the placebo estimation confirm the validity of our approach[8].

[7]They are available from the author upon request.

[8]If the common trend assumption were not respected would need to resort to, for example, the estimation technique proposed by Abadie (2005).

## 2.2 The machine learning algorithm

Random forests, introduced by Breiman (2001), are a widely used algorithm for statistical learning. The asymptotic behavior of quantities estimated via random forests has been widely studied (see e.g. Wager and Walther (2015), Athey et al. (2019)). Random forests based on the "aggregation" of random trees. The basic idea of a random tree is to divide the predictor space into distinct and non-overlapping regions. For every observation that falls into a given region, the same prediction is then made. An example of random tree is presented to reader for completeness in Figure 1. One drawback of the random tree is that it suffers from high variance. Random forests aim at reducing the variance of the trees, by decorrelating them and then aggregating their predictions (see Wager and Athey (2018)). In the present work, we rely on the method proposed by Athey et al. (2019), which has been implemented in the R package "grf". For simplicity, we present the estimation algorithm in two steps. First, we show how heterogeneity is tackled using a single regression tree, in the spirit of Breiman et al. (1984) with a modified splitting criterion borrowed from Wager and Athey (2018). Next, in order to reduce the usually high variance characterizing a single tree, the approach is extended to generalized random forests in the spirit of Athey et al. (2019) and Gulyas and Pytka (2020). The tree-based procedure consists in partitioning the dataset into smaller subsamples in which individuals exhibit similar treatment effects (i.e. log hourly wage changes) and at the same time the differences in wage changes between subsamples are maximized.

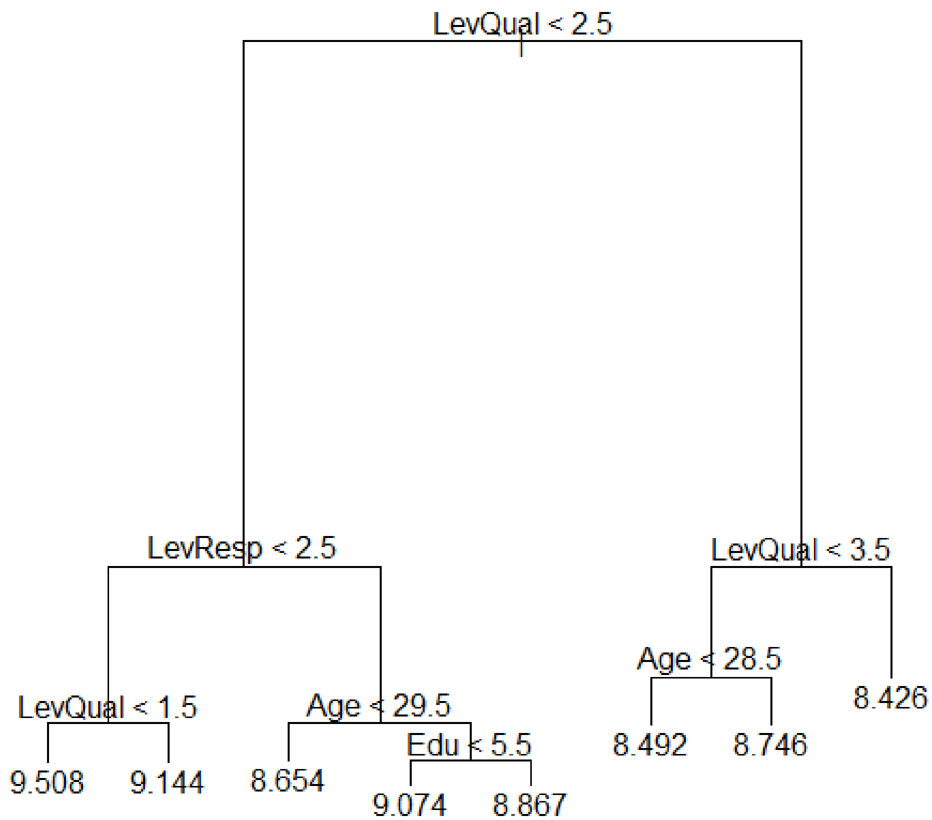


Figure 1: An example of random tree. In this Figure, a single random tree is represented as an example. The outcome variable is log real monthly wage and the explanatory variables those presented in Section 2.1. This tree has been estimated with the optimal values of the tuning parameters reported in Table 1. Treatment $W$ is here binary, e.g. treated observations are those closer than 15 minutes travel time from the border. Age stands for individual age, Edu is the level of educational qualification, where 9 is the highest (university degree) and 1 is the lowest (compulsory education only), LevResp stands for the level of responsibility and ranges from 1 (top management) to 5 (no management responsibility). LevQual refers to the level of skills needed to perform the job and ranges from 1 (the most complex tasks, requiring creativity) and 4 (i.e. only simple and repetitive tasks).

The procedure of building a tree can be characterized in the recursive way by the following algorithm, in which in each data partition (called also a node or a leaf) the treatment effect is estimated from equation (2) separately:

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

2. For each partitioning variable $X^k$ and its every occurring value $x$, split partition $\mathcal{S}$ into two complementary sets of individuals $i$ such that $\mathcal{S}_1 = \{i \in \mathcal{S} : X_i^k \leq x\}$ and $\mathcal{S}_2$ is what is left of $\mathcal{S}$ after eliminating $\mathcal{S}_1$ and estimate wage changes (i.e. treatment effects) $\gamma_{2,1}$ and $\gamma_{2,2}$ for both partitions by running two separate regressions of equation (2) on $\mathcal{S}_1$ and $\mathcal{S}_2$.

3. Chose the variable $X^k$ and the value $x$ that maximizes

$$(\gamma_{2,2} - \gamma_{2,1})^2 \frac{n_1 n_2}{N^2}, \tag{3}$$

   where $n_1$ and $n_2$ are the sizes of the subsamples $\mathcal{S}_1$ and $\mathcal{S}_2$, respectively, and N is the sample size of $\mathcal{S}$.

4. If the quantity computed in (3) is smaller than a certain tolerance improvement threshold, then the algorithm stops. Otherwise, it goes back to step 2 and repeats the splitting procedure for $\mathcal{S}_1$ and $\mathcal{S}_2$ separately.

We use a modified splitting criterion with respect to Breiman et al. (1984). Indeed, following Wager and Athey (2018), in eq. (3) we maximize between-group differences of wage changes, instead of building the tree minimizing the squared sum of residuals.

### 2.2.1 Generalization to random forests

Tree-based models have an intuitive graphical illustration. However, in comparison to other machine-learning methods their prediction accuracy is relatively low (Mullainathan and Spiess (2017)). Tree-based methods usually perform very well against other ML algorithms with classification tasks, where the outcome variable is binary or categorical. However, the outcome variable in this paper is continuous. Random forests proposed by Breiman (2001) are a refinement of the baseline method that addresses the typical concerns of tree-based models. The general idea behind random forests 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. The final prediction of a random forest is the average of individual predictions of many trees from the forest. Further, 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 heterogeneous treatment effects were estimated in those partitions. No data was used for both tasks (Athey et al. (2019), Gulyas and Pytka (2020))[9].

### 2.2.2 Implementation details

We apply the generalized random forest algorithm to a diff-in-diff framework, as it has been done by Gulyas and Pytka (2020)[10]. We use the orthogonalized version of the generalized random forest algorithm, which is meant to offer doubly robust estimates of the ATE (in the sense of Athey et al. (2017)). The variances of the ATE and the CATE are computed via the bootstrap of little bags as proposed by Sexton and Laake (2009). The results obtained with causal forest may be sensitive to the choice of the tuning parameters (Athey et al. (2017)). In order to alleviate this issue, all the estimations presented in the present paper have been performed using the "optimal" values of the tuning parameters. These optimal values have been obtained via cross-validation and are reported in Table 1.

Cross-validation is performed as follows: first, a number of random points in the space of possible parameter values is drawn. In our application, 100 distinct sets of parameter values are chosen (this is the default of the grf package). For each set of parameter values, a forest is trained and the out-of-bag error is computed. This error measure corresponds to the one developed by Nie and Wager (2021). Given these error estimates for each set of parameters, finally a smoothing function is applied, to determine the optimal parameter values. The optimal

[9]One implicit assumption of the machine learning algorithm used here is that individuals are all independent from one another. This could be violated essentially in two cases: (i) the flow of immigrants causes native workers to move or to modify their labour supply decisions. However, Beerli, Peri, Ruffner and Siegenthaler (2021) have already shown for Switzerland that there has been no crowding-out of natives due to the inflow of foreigners following the free movement reform. (ii) it could be that migrants create networks, i.e. future flows of migrants are more likely to arrive where there is already an existing community of foreigners. In the empirical analysis we will simply rule out this last possibility and assume that individuals are mutually independent. Indeed, we believe that the phenomena of immigrant network is far more relevant on a longer time horizon and in a large country like the US, rather than in the short term and in a small country like Switzerland.

[10]An alternative would have been to apply one of the strategies proposed by Chernozhukov et al. (2020) for inference on heterogeneous treatment effects.

parameters are the ones minimizing the predicted smoothed error on a new random draw of possible parameter values.

The number of trees for each random forest is fixed to 10'000. Standard errors are clustered at the level of the mobility regions (MS). Each cluster is given the same weight. In addition, sample weights are used in all estimations. The random forest is constructed with the honesty property (Wager and Athey (2018)), and the honesty fraction (i.e. the fraction of the data used to determine splits) is equal to 0.55.

Table 1: Optimal values of the tuning parameter of the forest, obtained via cross-validation.

| **Name of parameter** | **Meaning** | **Optimal value** |
|---|---|---|
| sample.fraction | Fraction of the data used to build each tree. | 0.4682 |
| mtry | Number of variables tried for each split. | 21 |
| min.node.size | A target for the minimum number of observations in each tree leaf. | 4 |
| honesty.fraction | The fraction of data that will be used for determining splits. | 0.5553 |
| honesty.prune.leaves | If equal to 1, prunes the estimation sample tree such that no leaves are empty. If equal to 0, keep the same tree as determined by the splits. | 0 |
| alpha: | A tuning parameter that controls the maximum imbalance of a split. | 0.0061 |
| imbalance.penalty: | A tuning parameter that controls how harshly imbalanced splits are penalized. | 0.4500 |

# 3 The data

The dataset used in the present paper is the Swiss Earnings Structure Survey (SESS) which collected demographic and labor market information every two years starting in 1994. The survey constitutes a repeated cross-section representative of all individuals working in Switzerland. It includes Swiss individuals as well as foreigners living and working in Switzerland and cross-border commuters. It includes detailed information about workers, their demographic characteristics, and their place of work identified to the level of commuting zones (called MS regions)[11].

[11] Of course, it would be interesting to have information not only on the commuting zone (MS region), but also on the municipality in which each individual actually works. However, the Swiss data protection law has changed since BRSP 2021 performed their study; hence, it is not anymore possible to obtain data with geographical information at such a fine level. In Section 4, we show that we are nevertheless able to replicate BRSP 2021 results.

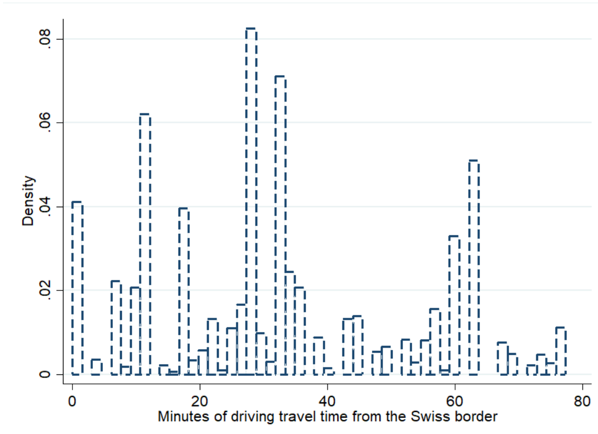


Figure 2: Histogram of the distance to the Swiss border (i.e. closest border crossing), measured in driving travel time, of each observation in our sample in year 2010. The exact location of each observation has been approximated by the MS region in which he/she works. Top 5% of travel time has been trimmed for readability.

We measure distance from the border in terms of travel time by car, as it is the most prevalent mean of transportation used by commuters employed in Switzerland[12]. Figure 2 shows the histogram of driving travel time to the closest Swiss border crossing in year 2010. From this Figure we deduce that in our dataset there is a large variation in treatment intensity, with most of the observations ranging between zero and 80 minutes of travel time to the border.

Figure 3 shows the evolution of wage over time (from 1994 to 2008), by type of work permit. From Figure 3 we deduce that changes in wage density have been almost negligible for Swiss nationals, as well as for holders of B and C permits. On the other hand, the mass of the density slightly shifted to the right for holders of G and L permits, thus indicating an increase in their average wage.

This finding provides support to the hypothesis that the migration inflows were not constituted by "cheap labor force" that caused a decrease in natives' wages. Indeed, in our sample, the percentage of holders of a university degree is equal to around 6.7% among Swiss nationals, whereas it is 7.7% among holders of a "C" permit and more than 9.5% among holders of a "G" permit (i.e. cross-border commuters). This confirms that immigrant flows in Switzerland after the free movement reform have been mostly constituted by high-skilled workers. This descriptive finding is fully consistent with the analysis carried out by BRSP 2021.

[12]The share of commuters using a car was equal to 52% in 2017. Source: Federal Statistical Office data on Commuting (PEND).

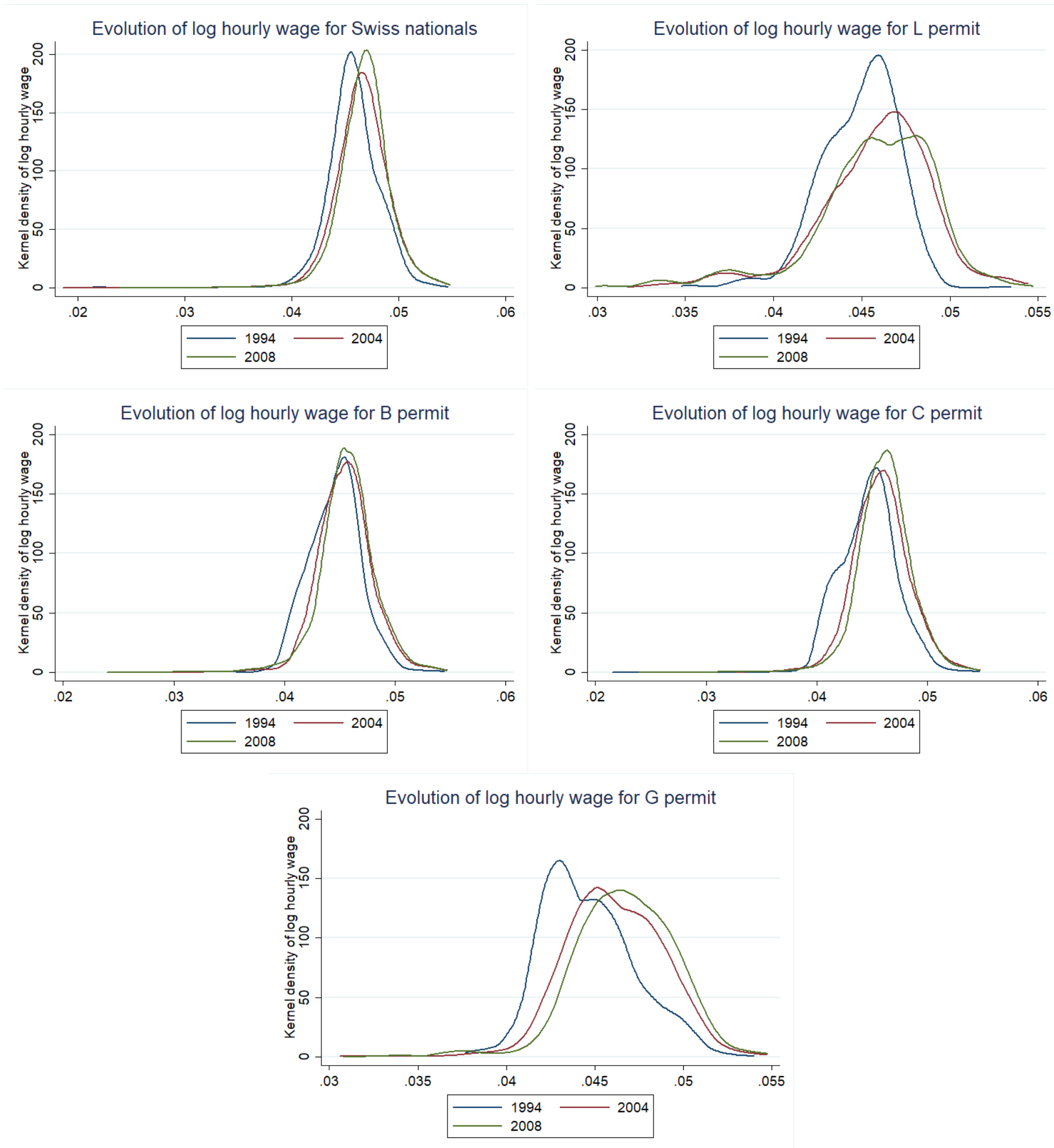


Figure 3: Kernel density estimation of log hourly real wages in 1994, 2004 and 2008, by type of work permit. Note that holders of the "C" work permit are considered as permanent residents.

The sample includes individuals between 18 and 65 years old, working in the private sector, with non-missing information for wages and hours worked. In particular, the data set contains the gross monthly wage for each individual worker (in the month of October) in Swiss Francs. This measure includes social transfers, bonuses, and one-twelfth of additional yearly payments.

We divide this measure by the number of hours worked in October, and use the consumer price index to deflate it into the real hourly wage of an individual worker at 2010 constant prices. We express hours worked as a fraction of the number of hours worked by a full-time worker, so that one unit is a full time equivalent (FTE). Missing values in the covariates are flagged and handled directly by the grf algorithm in the estimation part (Athey et al. (2019) and Robins and Rotnitzky (1995)). Workers are assigned to MS regions on the basis of the place of work (which is not necessarily the same as their place of residence). Both Swiss nationals, foreigners with different type of work permit (i.e. L, B or C) and cross-border commuters (i.e. work permit G) are included in our dataset.

Data on the minimum distance (in minutes of driving travel time) of each mobility zone (MS region) from the national border have been obtained via the Google API Console. Since driving is the most common commuting mode in Switzerland, we considered this as the transportation mode for the computation of all the commuting times. As mentioned in Section 2.1, in our causal forest estimations, we use as regressors or features the following variables: age, gender, education, degree of responsibility, type of contract, length of service, and level of qualifications required for the job. Regional labor market conditions are summarized by mobility region (MS) fixed effects. On the other hand, firm characteristics are proxied by the type of activity and by the number of employees[13]. A potential concern is that the industry composition of the mobility regions which are closer to the Swiss border may be relevantly different from that of the municipalities which are farther away from it. In order to tackle this issue, we follow the approach adopted by BRSP 2021, i.e. we introduce among the regressors the Bartik index, which is a weighted sum of industry employment growth rates. This index accounts for sector-driven demand trends that could affect regions differently due to their pre-existing industrial structure (Bartik (1991)). Our chosen set of explanatory variables closely resembles that used by BRSP 2021 in their analysis. As in their case, we claim that the exogeneity assumption is respected.

[13]Education is defined as: 1 = university, 2 = professional university, 3 = higher professional education, specialized school, 4 = teaching degree, 5 = high school, 6 = completed apprenticeship, 7 = professional training only acquired within the firm, 8 = compulsory education only, 9 = other type of education. Degree of responsibility is defined as: 1 = top manager, 2 = middle manager, 3 = lower manager, 4 = with some degree of responsibility, 5 = with no management function. Type of activity corresponds to one of 24 main sectors in which the NOGA classifications have been aggregated in the dataset. Type of job correspond to the skill requirements for the job and are defined as follows: 1 = the most difficult and demanding tasks, 2 = tasks requiring independent and highly qualified work, 3 = tasks requiring specialized professional knowledge, 4 = tasks requiring repetitive activities.

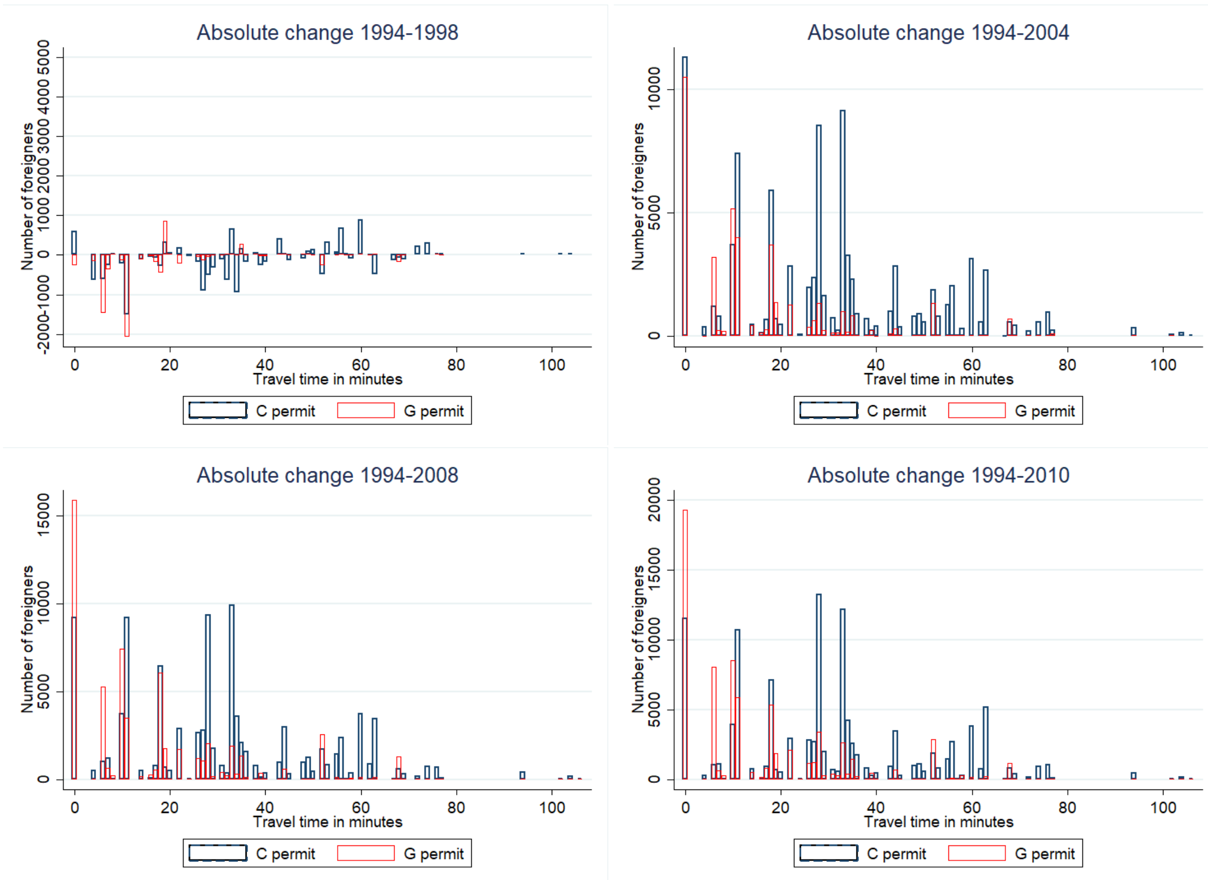


Figure 4: Change in the absolute number of cross border commuters and C permit holders, by minutes of travel time from the national border, over different time horizons. Sample weights are used in all the panels.

BPRS 2021 already shown that distance from the national border is a reasonable proxy for treatment intensity. Indeed, they show that most of the increase in both migrant and cross-border workers happened within 20 minutes of travel time from the border. In Figure 4 we confirm their finding.

Table 2: Summary statistics, full sample 1994-2010

| Variable | Mean | Std. Dev. | Min. | Max. | N |
|---|---|---|---|---|---|
| Age | 41.0068 | 11.7045 | 16 | 65 | 4’354’213 |
| Female dummy | 0.4898 | 0.4999 | 0 | 1 | 4’354’213 |
| Log gross monthly wage | 8.7288 | 0.4899 | 0 | 14.5621 | 4’354’213 |
| Foreigner dummy | 0.334 | 0.4716 | 0 | 1 | 4’354’213 |

In Table 2 we report descriptive statistics for the main variables of interest on the full sample

(1994-2010). The sample size is huge, with more than 4 millions individual-year observations. This also speaks in favour of our proposed machine learning approach, which is known for being specifically fit for dealing with large datasets. In Tables 6-8, reported in Appendix B, we report summary descriptive statistics for the main variables, for the three phases under scrutiny (pre-liberalization (1994-1998), transition (1998-2004) and full liberalization (2004-2016)) and for four distance bins from the Swiss border (i.e. less than 10, between 10 and 20, between 20 and 30, more than 30 minutes driving travel time). These Tables show that the mean values of these variables are rather close in the different distance bins. We use these tables as an (informal) balance check before estimating the CATEs.

# 4 Results and discussion

In this Section our main causal forest results are presented. A series of both economic and machine learning robustness checks are also reported and discussed. As a first step, we estimate the importance of each explanatory variable in the construction of the forest.

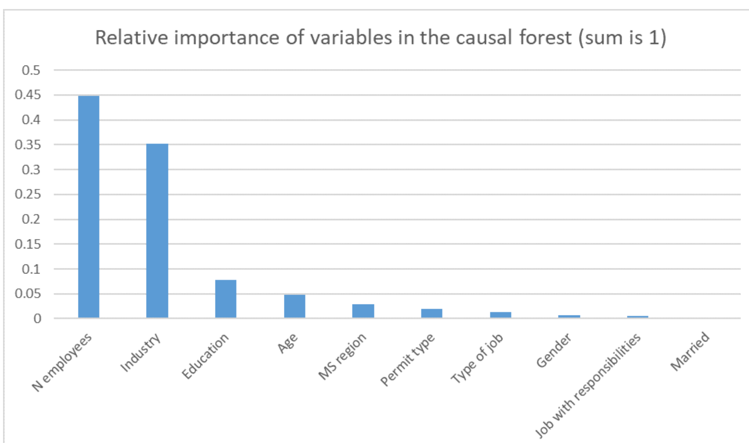


Figure 5: Graph of variable importance. In this Figure, the relative importance of the most important explanatory variables used to construct the causal forest is reported. This quantity is a weighted sum of how many times a certain feature was split on at each depth in the forest (Wager and Athey (2018)). The outcome variable is log real hourly wage and the explanatory variables are those presented in Section 2.1.

A summary of variable importance is reported in Figure 5. From this Figure, we deduce that the most relevant features, i.e. the variables which determine more splits in the construction of the forest, are the number of employees of the firm, the type of activity (industry), the level of education, and age. For this reason, in the following, we present the estimated CATEs for all the possible values of these four covariates.

## 4.1 Baseline causal forest results

As mentioned in Section 2, in our baseline results we define treated observations as those closer than 15 minutes of driving time to the closest national border crossing. In this we follow BRSP 2021, who claim that most of the impact of opening the borders took place within this threshold and define highly treated regions as those closer than 15 minutes to the border. We take as start date of the treatment year 2004.

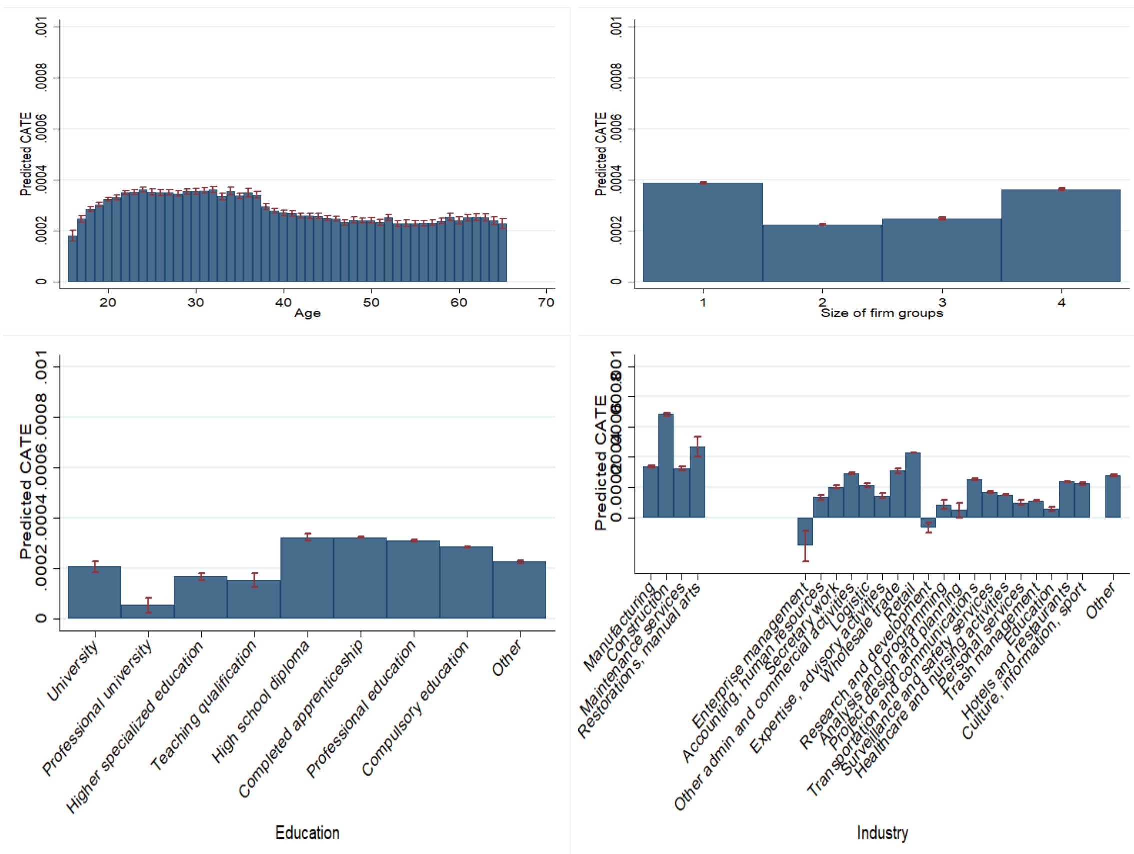


Figure 6: Predicted CATEs for each possible value of the four most important covariates (age, number of employees, type of activity, education). All other covariates are kept constant at their mean value in the sample (or median value for the discrete ones). Note that type of activity has been renamed industry for simplicity. We divide firms in four groups, according to the number of employees: (i) less than 20, (ii), between 20 and 200, (iii) between 200 and 2000, (iv) more than 2000. The red bars stand for the 95% confidence intervals, which have been obtained via bootstrapping.

In Figure 6 we predict treatment effect for each possible value of the four most important covariates, while keeping all the other covariates constant at their mean value in the sample (or median value for the discrete ones). From Figure 6, we deduce that treatment effect has been positive especially for young workers (i.e. in their 20s and 30s up to around 38), as well as for individuals working either in very small firms (i.e. less than 20 people) or very large ones (i.e. more than 2000 employees). Further, the predicted treatment effect is positive for workers with any educational qualification, except for those with a degree from a professional university (in this last case the treatment effect is not statistically significant from zero). Moreover, the predicted CATE is also positive for those working in the construction industry, in manual arts and in wholesale trade. However, even when the predicted CATE is statistically significant from zero, it is quantitatively rather small. For example, a worker in his 20s would witness an increase in his log wage equal to around 0.00035 log points. This means that, for example, an individual with a gross monthly wage of 4500 Swiss francs would see her wage increase by around two francs per month. A quantitatively similar impact is found for workers in very small or in very large firms. Is worth noting that, differently from BRSP 2021, who found a positive treatment effect for high-educated workers only, we find a positive treatment effect for all educational levels, possibly with the only exception of degrees from a professional university. Indeed, as represented in Figure 5, the size of the firm and the industry of employment have a notably higher relevance than education in determining the heterogeneous treatment effect.

| | Estimate | Standard error | P-value |
|---|---|---|---|
| Mean forest prediction | 1.01200 | 0.10426 | 2.2e-16 |
| Differential forest prediction | 3.27409 | 0.68541 | 8.904e-07 |

Table 3: Calibration test for the estimated causal forest. The test calibration of the forest computes the best linear fit of the target estimand using the forest prediction (on held-out data) as well as the mean forest prediction as the sole two regressors. A coefficient of 1 for ‘mean.forest.prediction‘ suggests that the mean forest prediction is correct. Further, if the coefficient for ‘differential.forest.prediction‘ is significantly greater than 0, then we can reject the null of no heterogeneity.

In order to assess whether heterogeneity in treatment effect is statistically significant or if it is just noise, we perform what Athey at al. (2019) call an omnibus evaluation of the quality of the random forest estimates via calibration. This test calibration of the forest explicitly provides information on the presence of heterogeneity. In Table 3, we report the result of this test for our estimated causal forest. The test calibration of the forest computes the best linear fit of the target estimand using the forest prediction (on held-out data) as well as the mean forest prediction as the sole two regressors. A coefficient of 1 for ‘mean.forest.prediction‘ suggests that the mean forest prediction is correct. In our case, the value is notably close to 1 (i.e. 1.01), thus suggesting that our estimated forest has a good fit to the data. Further, the p-value of the ‘differential.forest.prediction‘ coefficient also acts as an omnibus test for the presence of heterogeneity: if the coefficient is significantly greater than 0, then we can reject the null of no heterogeneity. This is what happens in our case: the estimated coefficient for the differential forest prediction is statistically different from zero, thus confirming the presence of heterogeneity in treatment effect. All in all, we find evidence of heterogeneous treatment effect. This is consistent with BRSP 2021, who find a positive impact for high-educated native workers and with Dustmann et al. (2017), who claim that the impact of migration flows on natives' labor market outcomes can greatly vary across the wage distribution. Nevertheless, all of our estimated CATEs are rather tiny, ranging approximately between 0.0002 and 0.00035 log points, hence it is not clear whether, from a policy-making perspective, we should worry about heterogeneity in treatment effects at all. However, given that the large majority of the estimated CATEs are positive, we agree with BRSP 2021 that, in the context of the Swiss migration reform, native and foreigner workers are strategic complements rather than strategic substitutes.

## 4.2 Robustness checks

In this Section, we present the results of a number of robustness checks. First, one could argue that the Agreement on the Free Movement of Persons was not the only policy change entering into force in 2002. Indeed, it was part of the so-called "package of Bilateral agreements I", which included agreements on a variety of topics, such as transportation and trade. It is likely that, e.g. trade liberalizations also affected regions close to the border more than regions further away. However, BRSP 2021 have already shown (with the same dataset and in the same time frame) that the mean effect that they identify on wages is not driven by this simultaneous policy change. Second, the above-mentioned free movement agreement also allowed Swiss individuals to go working in the border area of the neighbouring countries. However, we do not consider this effect because, as explained by BRSP 2021, the rate of employment of Swiss individuals in the border regions of Italy, France and Germany did not significantly change in the years following the Bilateral Agreements I. Indeed, the wage differential was always strongly in favor of the Swiss side of the border. Third, another concern about the causal interpretation of the results obtained is that resident workers may respond to immigrant inflows by leaving the labor market or the region affected by the reform (Borjas (2003), Dustmann et al. (2017)). BRSP 2021 already performed this robustness check with reference to the same time horizon considered in the present paper and found no evidence that the greater availability of cross-border workers affected the flow of residents between treated and control regions. Further, they find no evidence of an impact of opening the borders on regional population size (BRSP 2021). In the following, we present the results of two additional checks. In the first one, we define treated observations as those closer than 20 (instead than 15) minutes driving time from the Swiss border. In the second one, we define the start date of the treatment as 2002 (instead of 2004 as in our baseline results), e.g. when the Free Movement Agreement entered into force but quotas and priority requirement were still in place. Finally, we estimate CATEs for the placebo period (i.e. 1994-1998) to check that the identifying assumptions of our model are respected.

#### 4.2.1 Treated observations are those closer than 20 minutes to the border

When treatment group is defined as all those observations which are closer than 20 minutes of travel time from the closest border crossing, our main results are confirmed. From Figure 7 we deduce that treatment effect is positive for young individuals (i.e. in their 20s up to around 38) and for those working in either very small (i.e. less than 20 employees) or in very large firms (i.e. more than 2000 employees). On the other hand, we find that the predicted treatment effect is positive and almost equal for people working in all of the industries considered. As far as education is considered, we record the highest positive CATEs for individuals with either completed apprenticeship or professional education, whereas the CATE is essentially zero for workers with a university or a professional university degree. From Table 4, we find evidence that the fit of the causal forest with this alternative definition of binary treatment (i.e. being closer than 20, instead than 15, to the Swiss border) is rather good. Indeed, the coefficient for the mean forest prediction is close to 1 (i.e. 0.99). Further, the coefficient for the differential forest prediction is statistically different from zero, thus confirming that treatment effect is heterogeneous.

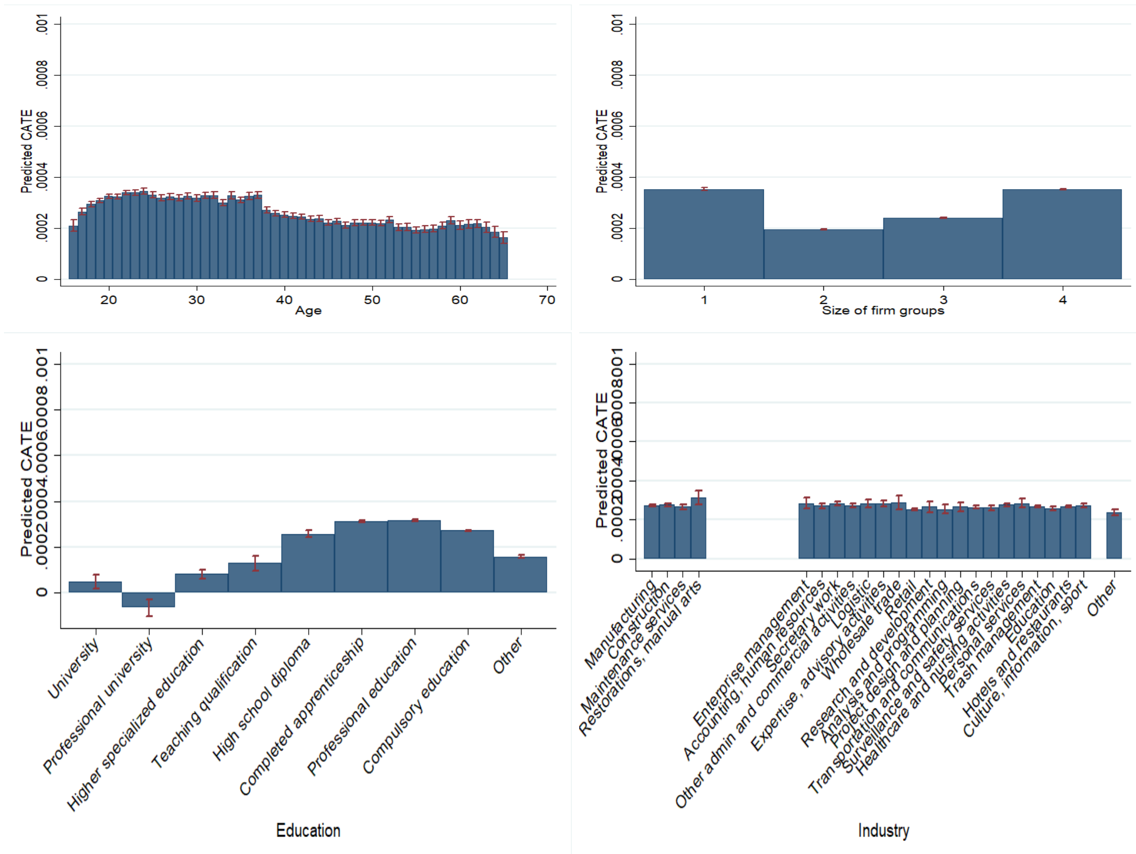


Figure 7: Predicted CATE for each possible value of the four most important covariates (age, number of employees, type of activity, education). All other covariates are kept constant at their mean value in the sample (or median value for the discrete ones). Note that type of activity has been renamed industry for simplicity. We divide firms in four groups, according to the number of employees: (i) less than 20, (ii), between 20 and 200, (iii) between 200 and 2000, (iv) more than 2000. The red bars stand for the 95% confidence intervals, which have been obtained via bootstrapping.

| | Estimate | Standard error | P-value |
|---|---|---|---|
| Mean forest prediction | 0.99604 | 0.11521 | 2.2e-16 |
| Differential forest prediction | 4.23852 | 0.79288 | 4.504e-08 |

Table 4: Calibration test for the estimated causal forest. The test calibration of the forest computes the best linear fit of the target estimand using the forest prediction (on held-out data) as well as the mean forest prediction as the sole two regressors. A coefficient of 1 for ‘mean.forest.prediction‘ suggests that the mean forest prediction is correct. Further, if the coefficient for ‘differential.forest.prediction‘ is significantly greater than 0, then we can reject the null of no heterogeneity.

#### 4.2.2 Start date of the treatment is 2002

As a further robustness check, in this subsection we consider as start date of the treatment year 2002 instead of 2004. Our baseline results are confirmed, in the particular we still find the highest positive CATES for young workers (i.e. in their 20s) and for those working in large firms (i.e. with more than 2000 employees). Further, we find that the estimated treatment effect is positive for workers with all educational level, but in particular for those with high school diploma or completed professional apprenticeship. However, the difference in CATEs between those workers and those with a university degree is notably small, i.e. around 0.0001 log points, which corresponds to a difference in treatment effect of around 0.5 Swiss francs per month for an individual earning 5000 Swiss francs monthly. Hence, we deduce that these differences are too small to be of interest for policy-making. Finally, we find that the CATEs is especially positive for those working in enterprise management and education. Nevertheless, like in the previous section, all the estimated CATEs are tiny, ranging from zero to around 0.00035 log points. The actual impact on wages is hence negligible in all the cases presented. When looking at Table 5 we find that the model fit is slightly worse than our two previous specifications (i.e. 0.97 vs 0.99 and 1.01, where the best fit would be 1). Hence, we deem that the predictions obtained in this section are slightly less accurate than those presented under the baseline specification, as well as under the first robustness check performed.

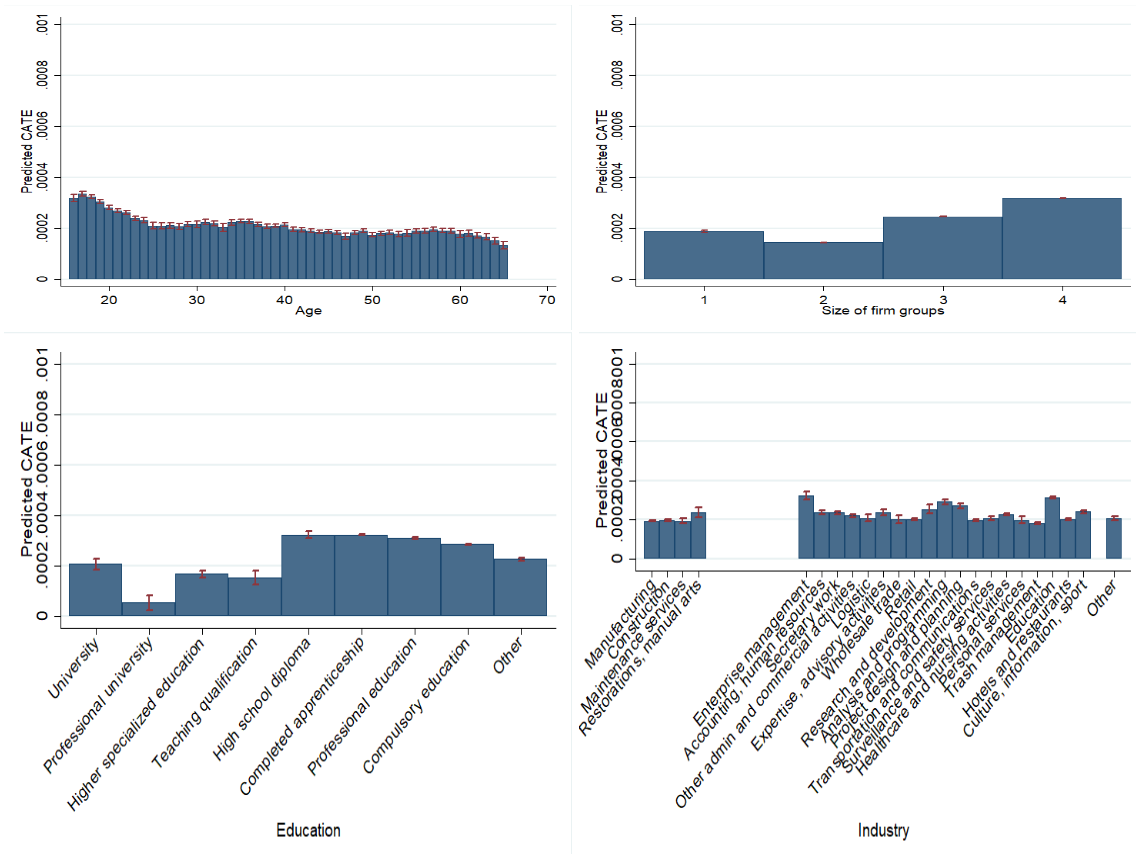


Figure 8: Predicted CATE for each possible value of the four most important covariates (age, number of employees, type of activity, education). All other covariates are kept constant at their mean value in the sample (or median value for the discrete ones). Note that type of activity has been renamed industry for simplicity. We divide firms in four groups, according to the number of employees: (i) less than 20, (ii), between 20 and 200, (iii) between 200 and 2000, (iv) more than 2000. The red bars stand for the 95% confidence intervals, which have been obtained via bootstrapping.

| | Estimate | Standard error | P-value |
|---|---|---|---|
| Mean forest prediction | 0.971610 | 0.096976 | 2.2e-16 |
| Differential forest prediction | 3.654213 | 0.790626 | 1.901e-06 |

Table 5: Calibration test for the estimated causal forest. The test calibration of the forest computes the best linear fit of the target estimand using the forest prediction (on held-out data) as well as the mean forest prediction as the sole two regressors. A coefficient of 1 for ‘mean.forest.prediction‘ suggests that the mean forest prediction is correct. Further, if the coefficient for ‘differential.forest.prediction‘ is significantly greater than 0, then we can reject the null of no heterogeneity.

### 4.2.3 Estimation on placebo years (1994-1998)

In this subsection, we estimate the same causal forest of Section 4.1 on the placebo years, i.e. 1994-1998. As in Section 4.1, treatment group is defined here as all the observations which are closer than 15 minutes driving time from the border and the start date of the treatment is 2004.

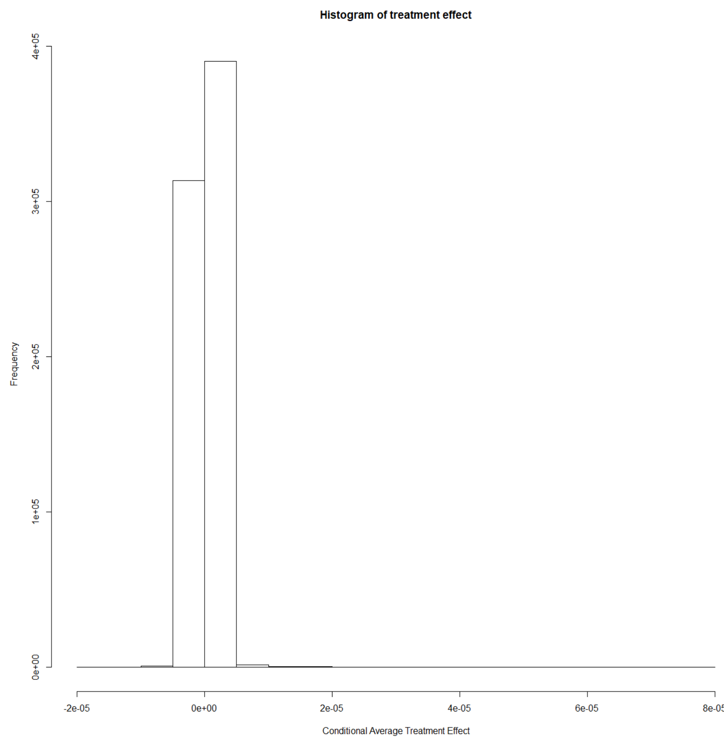


Figure 9: Histogram of Conditional Average Treatment Effect in the placebo years

It could be that the industry structure and the labor market is different in treated areas than in non-treated areas - and with different time trends. This could induce spurious results. By using exactly the same data source and considering an identical time frame, BRSP 2021 have already shown that there was no systematic pre-trend in wages before the introduction of the Free Movement Agreement, and hence the diff-in-diff analysis can be validly performed. Our results (Figure 9) show that the estimated CATEs are essentially zero in the placebo period.

# 5 Conclusion

In the present paper we study the heterogeneous impact of opening the borders on wages, for different groups of workers. Our results are complementary to those of BRSP 2021, as we find confirmation of their main findings on the statistically insignificant average impact of the reform on wages. On the other hand, we find that this impact has been heterogeneous across individual characteristics. In particular, we find evidence of a positive impact of the free movement reform for young individuals and for those working for firms with a large number of employees. Also, we find a statistically significant positive treatment effect for workers in the construction and in the retail industry.

By applying causal forest methods in the spirit of Athey et al. (2019), we compute the conditional average treatment effects (CATEs) for different groups of workers, in order to assess which individuals have been most affected by this regional policy change. The present paper is one of the first to introduce supervised machine learning methods in order to answer an economic research question. The advantage of choosing this method (i.e. causal forests) lies in the notably higher degree of heterogeneity in treatment effect which we can allow for. However, while we find that heterogeneity in treatment effect is statistically significant in this framework, the size of the estimated differences in CATEs is notably small, being on average around 0.0001 log points. This corresponds to differences in the monetary impact of the reform equal to around 0-5-1 Swiss franc per month, not enough to be relevant in the light of policy-making.

## References


1. Abadie, A. (2005). Semiparametric difference-in-differences estimators. The Review of Economic Studies, 72(1), 1-19.

2. Akgündüz, Y. E., & Torun, H. (2020). Two and a half million Syrian refugees, tasks and capital intensity. Journal of Development Economics, 145, 102470.

3. Athey, S., Imbens, G., Pham, T., & Wager, S. (2017). Estimating average treatment effects: Supplementary analyses and remaining challenges. American Economic Review, 107(5), 278-81.

4. Athey, S., Tibshirani, J., & Wager, S. (2019). Generalized random forests. The Annals of Statistics, 47(2), 1148-1178.

5. Bartik, T. J. (1991). Who Benefits from State and Local Economic Development Policies?. Books from Upjohn Press.

6. Basten, C., & Siegenthaler, M. (2019). Do immigrants take or create residents' jobs? Evidence from free movement of workers in Switzerland. The Scandinavian Journal of Economics.

7. Beerli, A., Ruffner, J., Siegenthaler, M., & Peri, G. (2021) The Abolition of Immigration Restrictions and the Performance of Firms and Workers: Evidence from Switzerland. American Economic Review.

8. Bertrand, M., Crepon, B., Marguerie, A., & Premand, P. (2017). Contemporaneous and Post-Program Impacts of a Public Works Program: Evidence from Cote d'Ivoire. Working paper.

9. Bitler, M. P., Gelbach, J. B., & Hoynes, H. W. (2006). What mean impacts miss: Distributional effects of welfare reform experiments. American Economic Review, 96(4), 988-1012.

10. Borjas, G. J. (2003). The labor demand curve is downward sloping: Reexamining the impact of immigration on the labor market. The quarterly journal of economics, 118(4), 1335-1374.

11. Borjas, G. J., Grogger, J., & Hanson, G. H. (2008). Imperfect substitution between immigrants and natives: a reappraisal (No. w13887). National Bureau of Economic Research.

12. Bratsberg, B., & Raaum, O. (2012). Immigration and wages: Evidence from construction. The economic journal, 122(565), 1177-1205.

13. Breiman, L. (2001). Random forests. Machine learning, 45(1), 5-32.

14. Breiman, L., Friedman, J., Stone, C. J., & Olshen, R. A. (1984). Classification and regression trees. CRC press.

15. Card, D. (2009). Immigration and inequality. American Economic Review, 99(2), 1-21.

16. Card, D. (1990). The impact of the Mariel boatlift on the Miami labor market. ILR Review, 43(2), 245-257.

17. Card, D., & DiNardo, J. (2000). Do immigrant inflows lead to native outflows?. American Economic Review, 90(2), 360-367.

18. Chernozhukov, V., Demirer, M., Duflo, E., & Fernandez-Val, I. (2020). Generic machine learning inference on heterogenous treatment effects in randomized experiments (No. w24678). National Bureau of Economic Research.

19. Davis, J., & Heller, S. B. (2017). Using Causal Forests to Predict Treatment Heterogeneity: An Application to Summer Jobs. American Economic Review, 107(5), 546-50.

20. De Chaisemartin, C., & d'Haultfoeuille, X. (2020). Two-way fixed effects estimators with heterogeneous treatment effects. American Economic Review, 110(9), 2964-96.

21. Dustmann, C., Schonberg, U., & Stuhler, J. (2017). Labor supply shocks, native wages, and the adjustment of local employment. The Quarterly Journal of Economics, 132(1), 435-483.

22. Dustmann, C., Frattini, T., & Preston, I. P. (2012). The effect of immigration along the distribution of wages. Review of Economic Studies, 80(1), 145-173.

23. Favre, S. (2011). The impact of immigration on the wage distribution in Switzerland. Available at SSRN 1915067.

24. Foged, M., & Peri, G. (2016). Immigrants' effect on native workers: New analysis on longitudinal data. American Economic Journal: Applied Economics, 8(2), 1-34.

25. Friedberg, R. M. (2001). The impact of mass migration on the Israeli labor market. The Quarterly Journal of Economics, 116(4), 1373-1408.

26. Glitz, A. (2012). The labor market impact of immigration: A quasi-experiment exploiting immigrant location rules in Germany. Journal of Labor Economics, 30(1), 175-213.

27. Gulyas, A., & Pytka, K. (2020). Understanding the sources of earnings losses after job displacement: A machine-learning approach. Working paper. University of Bonn and University of Mannheim, Germany.

28. Kleinberg, J., Lakkaraju, H., Leskovec, J., Ludwig, J., & Mullainathan, S. (2018). Human decisions and machine predictions. The quarterly journal of economics, 133(1), 237-293.

29. Kugler, A., & Yuksel, M. (2008). Effects of low-skilled immigration on US natives: evidence from Hurricane Mitch (No. w14293). National Bureau of Economic Research.

30. Llull, J. (2017). Immigration, wages, and education: A labour market equilibrium structural model. The Review of Economic Studies, 85(3), 1852-1896.

31. Mitaritonna, C., Orefice, G., & Peri, G. (2017). Immigrants and firms' outcomes: Evidence from France. European Economic Review, 96, 62-82.

32. Mullainathan, S., & Spiess, J. (2017). Machine learning: an applied econometric approach. Journal of Economic Perspectives, 31(2), 87-106.

33. Nie, X., & Wager, S. (2021). Quasi-oracle estimation of heterogeneous treatment effects. Biometrika, 108(2), 299-319.

34. Ottaviano, G. I., & Peri, G. (2012). Rethinking the effect of immigration on wages. Journal of the European economic association, 10(1), 152-197.

35. Peri, G. (2016). Immigrants, productivity, and labor markets. The Journal of Economic Perspectives, 30(4), 3-29.

36. Peri, G., Shih, K., & Sparber, C. (2015). STEM workers, H-1B visas, and productivity in US cities. Journal of Labor Economics, 33(S1), S225-S255.

37. Peri, G., & Sparber, C. (2011). Highly educated immigrants and native occupational choice. Industrial Relations: a journal of economy and society, 50(3), 385-411.

38. Peri, G., & Sparber, C. (2009). Task specialization, immigration, and wages. American Economic Journal: Applied Economics, 1(3), 135-69.

39. Poulos, J., Albanese, A., Mercatanti, A., & Li, F. (2021). Retrospective causal inference via matrix completion, with an evaluation of the effect of European integration on cross-border employment. arXiv preprint arXiv:2106.00788.

40. Prantl, S., & Spitz-Oener, A. (2018). The Impact of Immigration on Competing Natives' Wages: Evidence from German Reunification. Review of Economics and Statistics, 1-46.

41. Robins, J. M., & Rotnitzky, A. (1995). Semiparametric efficiency in multivariate regression models with missing data. Journal of the American Statistical Association, 90(429), 122-129.

42. Sexton, J., & Laake, P. (2009). Standard errors for bagged and random forest estimators. Computational Statistics & Data Analysis, 53(3), 801-811.

43. Wager, S., & Athey, S. (2018). Estimation and inference of heterogeneous treatment effects using random forests. Journal of the American Statistical Association, 113(523), 1228-1242.

44. Wager, S., & Walther, G. (2015). Adaptive concentration of regression trees, with application to random forests. arXiv preprint arXiv:1503.06388.

# Appendix A. The Bilateral Agreements

The Bilateral Agreements I between Switzerland and the European Union were signed in June 1999 and came into force in 2002. They included seven liberalization agreements: free movement of persons, technical barriers to trade, agricultural products, overland transport, public procurement, scientific and technological cooperation. A second set of Bilateral Agreements was signed in October 2004 and approved in referendum in June 2005. The agreement on the free movement of persons between Switzerland and the European Union progressively granted full bilateral access to each other's labor markets. Even if this agreement had been signed in 1999, the details of the implementation procedure were not clear until the ratification process took place in all the EU member States and in Switzerland by means of a referendum; this process was completed in 2000. The Swiss Federal Government decided that full free workers mobility would have been implemented in a gradual fashion.

Before 1999, cross-border-workers could be hired only subject to the priority requirement, i.e. the firms had to show that it was not possible to find an equally qualified resident worker within a reasonable time frame. Of course, this administrative procedure imposed a non-negligible cost on firms willing to hire cross-border workers. Moreover, cross-border workers were not allowed to work outside the Swiss Border region. Note that the definition of the Swiss border region did not change with the introduction of the free movement agreement, nor did it change between 2002 and 2007, the year in which it was finally abolished. As far as the workers coming to live in Switzerland (i.e. immigrants) are concerned, they were subject to yearly national quotas, which were set by the federal government. They were also subject to the priority requirement as well. Between 2002 and 2004 there was a transitional phase, in which both the status of cross-border workers and that of immigrant workers was liberalized. In particular, the work permit of cross-border workers was no longer bound to a particular position and remained valid for five years, whereas previously it was valid for one year only. Further, those workers were no longer required to have resided in the border region of the neighboring country for at least the previous six months. Moreover, they were allowed to commute weekly rather than daily in order to return to their home country. On 1st June 2004 the full liberalization phase started, meaning that both quotas (for immigrants) and the priority requirement (for both immigrants and cross-border workers) were eliminated.

# Appendix B: Additional data analysis

Table 6: Summary statistics in the pre-reform phase (1994-1998).

| **0-10 min distance bin** | **Mean** | **Std. Dev.** | **Min.** | **Max.** |
|---|---|---|---|---|
| Age | 37.7341 | 12.0701 | 16 | 65 |
| Female dummy | 0.6794 | 0.4667 | 0 | 1 |
| Log gross monthly wage | 8.1684 | 0.3069 | 7.1349 | 10.0487 |
| Married dummy | 0.6422 | 0.4794 | 0 | 1 |
| % foreigners (i.e. L, B, C, G permit) | 0.5368 | 0.4987 | 0 | 1 |
| **10-20 min distance bin** | **Mean** | **Std. Dev.** | **Min.** | **Max.** |
| Age | 39.721 | 12.3875 | 16 | 65 |
| Female dummy | 0.7761 | 0.4169 | 0 | 1 |
| Log gross monthly wage | 8.2093 | 0.2504 | 7.0353 | 10.0227 |
| Married dummy | 0.7011 | 0.4578 | 0 | 1 |
| % foreigners (i.e. L, B, C, G permit) | 0.3705 | 0.483 | 0 | 1 |
| **20-30 min distance bin** | **Mean** | **Std. Dev.** | **Min.** | **Max.** |
| Age | 39.524 | 12.5177 | 16 | 65 |
| Female dummy | 0.7412 | 0.438 | 0 | 1 |
| Log gross monthly wage | 8.2902 | 0.2884 | 7.3524 | 10.6767 |
| Married dummy | 0.6583 | 0.4743 | 0 | 1 |
| % foreigners (i.e. L, B, C, G permit) | 0.3393 | 0.4735 | 0 | 1 |
| **30+ min distance bin** | **Mean** | **Std. Dev.** | **Min.** | **Max.** |
| Age | 39.6275 | 12.3217 | 16 | 65 |
| Female dummy | 0.7347 | 0.4415 | 0 | 1 |
| Log gross monthly wage | 8.2408 | 0.2711 | 7.0699 | 10.5497 |
| Married dummy | 0.6797 | 0.4666 | 0 | 1 |
| % foreigners (i.e. L, B, C, G permit) | 0.2789 | 0.4484 | 0 | 1 |

Table 7: Summary statistics in the transition phase (1999-2004).

| **0-10 min distance bin** | **Mean** | **Std. Dev.** | **Min.** | **Max.** |
|---|---|---|---|---|
| Age | 38.0089 | 12.1509 | 16 | 65 |
| Female dummy | 0.5611 | 0.4963 | 0 | 1 |
| Log gross monthly wage | 8.3482 | 0.3409 | 7.3317 | 11.3601 |
| Married dummy | 0.5891 | 0.4920 | 0 | 1 |
| % foreigners (i.e. L, B, C, G permit) | 0.5589 | 0.4965 | 0 | 1 |
| **10-20 min distance bin** | **Mean** | **Std. Dev.** | **Min.** | **Max.** |
| Age | 39.7779 | 12.396 | 16 | 65 |
| Female dummy | 0.7218 | 0.4481 | 0 | 1 |
| Log gross monthly wage | 8.3684 | 0.3047 | 6.452 | 11.2161 |
| Married dummy | 0.6479 | 0.4776 | 0 | 1 |
| % foreigners (i.e. L, B, C, G permit) | 0.3773 | 0.4847 | 0 | 1 |
| **20-30 min distance bin** | **Mean** | **Std. Dev.** | **Min.** | **Max.** |
| Age | 39.4382 | 12.1705 | 16 | 65 |
| Female dummy | 0.6648 | 0.4721 | 0 | 1 |
| Log gross monthly wage | 8.4743 | 0.3824 | 6.8438 | 11.1618 |
| Married dummy | 0.6018 | 0.4895 | 0 | 1 |
| % foreigners (i.e. L, B, C, G permit) | 0.3164 | 0.4651 | 0 | 1 |
| **30+ min distance bin** | **Mean** | **Std. Dev.** | **Min.** | **Max.** |
| Age | 39.6899 | 12.3537 | 16 | 65 |
| Female | 0.6967 | 0.4597 | 0 | 1 |
| Log gross monthly wage | 8.381 | 0.3254 | 6.2086 | 11.1267 |
| Married dummy | 0.6538 | 0.4758 | 0 | 1 |
| % foreigners (i.e. L, B, C, G permit) | 0.2759 | 0.447 | 0 | 1 |

Table 8: Summary statistics in the full liberalization phase (2005-2016).

| **0-10 min distance bin** | **Mean** | **Std. Dev.** | **Min.** | **Max.** |
|---|---|---|---|---|
| Age | 40.8049 | 11.237 | 16 | 65 |
| Female dummy | 0.4434 | 0.4968 | 0 | 1 |
| Log gross monthly wage | 8.7826 | 0.5093 | 1.0986 | 14.3946 |
| Married dummy | 0.5247 | 0.4994 | 0 | 1 |
| % foreigners (i.e. L, B, C, G permit) | 0.5366 | 0.4987 | 0 | 1 |
| **10-20 min distance bin** | **Mean** | **Std. Dev.** | **Min.** | **Max.** |
| Age | 41.7437 | 11.4502 | 16 | 65 |
| Female dummy | 0.4687 | 0.499 | 0 | 1 |
| Log gross monthly wage | 8.7956 | 0.5065 | 0 | 14.2381 |
| Married dummy | 0.5402 | 0.4984 | 0 | 1 |
| % foreigners (i.e. L, B, C, G permit) | 0.43 | 0.4951 | 0 | 1 |
| **20-30 min distance bin** | **Mean** | **Std. Dev.** | **Min.** | **Max.** |
| Age | 40.6925 | 11.4652 | 16 | 65 |
| Female dummy | 0.4605 | 0.4984 | 0 | 1 |
| Log gross monthly wage | 8.8451 | 0.5375 | 0 | 14.5621 |
| Married dummy | 0.512 | 0.4999 | 0 | 1 |
| % foreigners (i.e. L, B, C, G permit) | 0.3217 | 0.4671 | 0 | 1 |
| **30+ min distance bin** | **Mean** | **Std. Dev.** | **Min.** | **Max.** |
| Age | 41.3069 | 11.7737 | 16 | 65 |
| Female dummy | 0.4684 | 0.499 | 0 | 1 |
| Log gross monthly wage | 8.7439 | 0.449 | 0 | 14.4906 |
| Married dummy | 0.5351 | 0.4988 | 0 | 1 |
| % foreigners (i.e. L, B, C, G permit) | 0.2698 | 0.4438 | 0 | 1 |