Employment and Social Development Canada
Emploi et Développement social Canada
Canada

A Practical Guide to Implement Modified Causal Forests in Estimating Causal Impacts of Labour Market Programs

Quantitative Impact Assessment (QIA) Workshop

February 16, 2026

Evaluation Directorate - Strategic and Service Policy Branch

Overview of the Presentation

Purpose: To present the step-by-step procedure of implementing Modified Causal Forest (MCF), a causal machine learning approach in the impact analysis of labour market program.

Overview of Impact Analysis

- ❑ Traditional Approach
- ❑ Causal Machine Learning

Methodological Framework

- ❑ 10 Steps to Implement MCF using Synthetic Data.

Context

Fundamental Problem of Causal Inference

- We cannot observe both potential outcomes for the same individual at the same time. In other words, cannot see what *would have happened* under the alternative treatment or condition once one outcome has already occurred.

Potential Outcome Framework

- It defines and estimates causal effects by addressing the **fundamental problem of causal inference.**
- The framework estimates **causal effects** as the difference between average* outcomes with and without treatment.

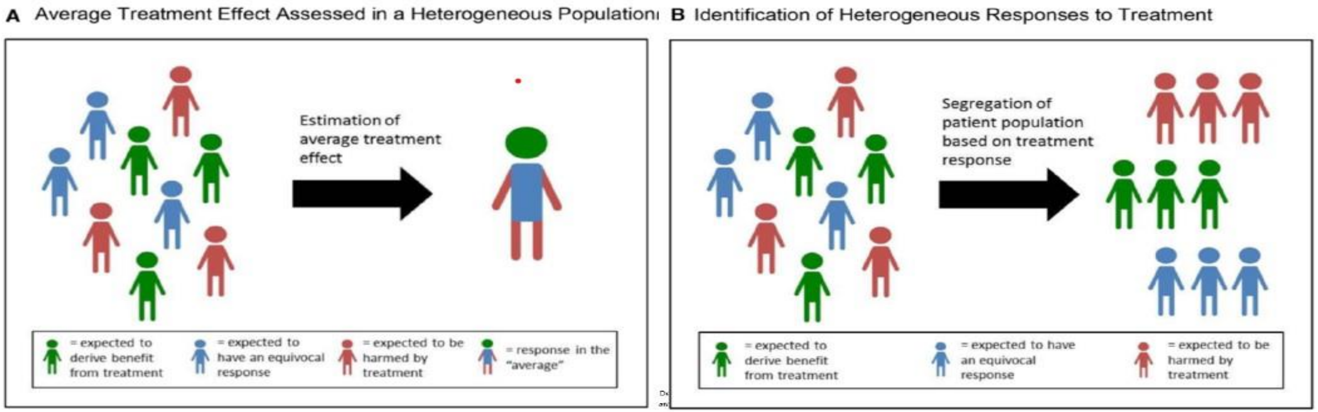
Typically, a **non-experimental design** is used to address the fundamental problem of causal inference in **observational studies**

- Estimate population or average causal effects by constructing a counterfactual scenario using control group whose observable characteristics are on average equivalent to the treated group.
- The well-established method to estimate causal effects is propensity score matching and difference-in-differences (**Traditional Methods** hence forth).

*Note: we cannot directly compute individual causal effects.

Context (cont’d)

- **Treatment effect heterogeneity** refers to the variation in the impact of the labour market programs across different socio demographic groups.
- Traditional methods are **not optimal** for identifying treatment effect heterogeneity.
- Causal Machine Learning adapts Machine Learning methods to answer well identified causal questions. Causal Forest (CF) is a tree based causal machine learning approach.
- Modified Causal Forest (MCF) introduced by Lechner (2019) can efficiently estimate treatment effect at most granular level, allowing for a more robust and data-driven analysis of effect heterogeneity and thereby also uncovering “**what works for whom**”.



Causal Forest (Basic Concepts)

Random forests can be transformed into **Causal Forests,** provided that the classical identification assumptions for causal effects are met and some modifications in algorithm* are made.

Identification Assumptions:

- ❑ Conditional Independence Assumption (CIA)
- ❑ Common Support (CS)
- ❑ Stable Unit Treatment Value Assumption (SUTVA)
- ❑ Exogeneity of Confounders

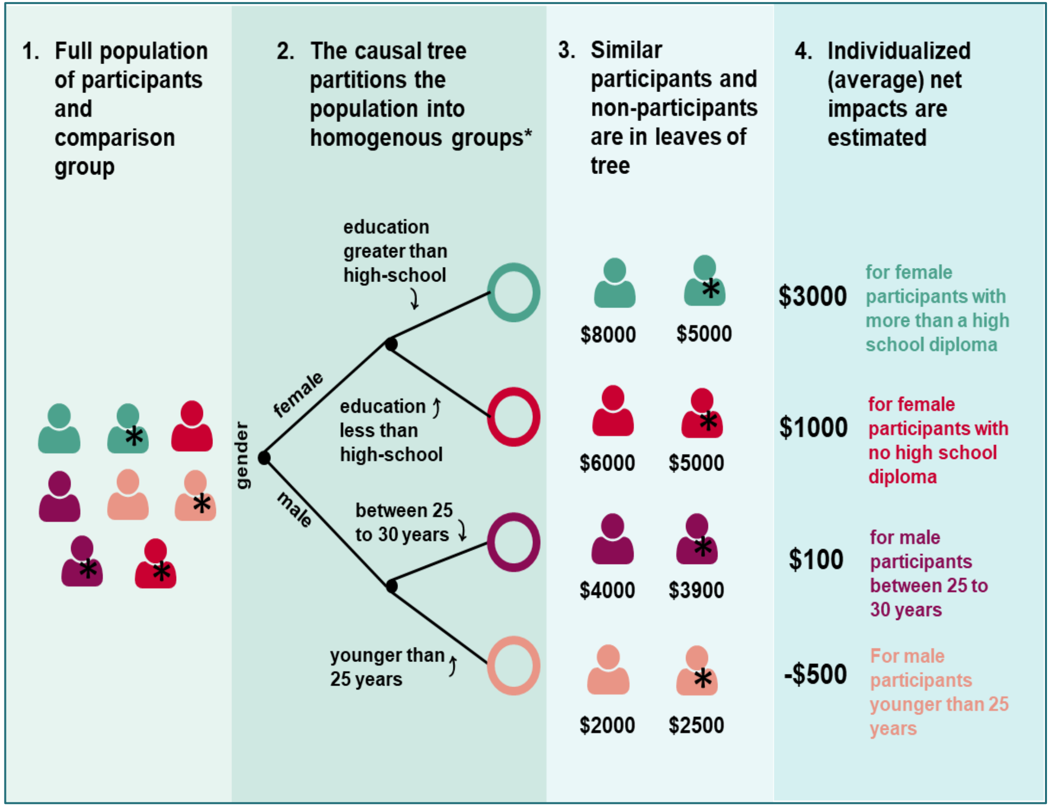
Sample Splitting in Causal Forest:

- ❑ The original sample is **splitted equally** into two disjoint sets known as training set (50%) and honest set (50%).
- ❑ **Training set** is used determine how to partition the feature space (i.e., how to grow the tree) whereas **honest set** is used to estimate treatment effects within each leaf. This prevents the same data from being used for **both model selection and effect estimation.**

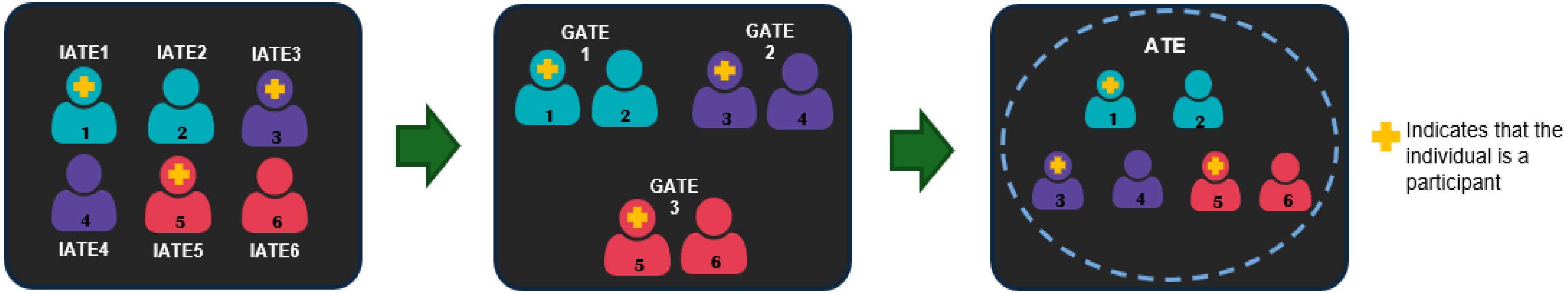
*The objective function of CF aims to maximize heterogeneity in treatment effects across leaves

Visualizing a Causal Tree

- Trees recursively split data into non-overlapping neighbourhoods.
- At each node, the split is chosen to satisfy an objective function and preserve a minimum number of participants and comparison cases.
- The **individualized (average) net impact** is estimated as the difference in mean outcomes between participants and non-participants in the leaves.



Key Parameters of Causal Forest



Individualized Average Treatment Effect (IATE)

Measures the average impact a treatment has for individuals with a given set of characteristics or covariate profile.

Represents causal treatment effects at the finest level of granularity

Group Average Treatment Effect (GATE)

By aggregating and weighting the IATEs over specific subgroups, the GATE can be estimated. Unlike traditional subgroup analyses, GATEs can be compared across groups.

Average Treatment Effect (ATE)

Represents the population average treatment effect.

Causal Forest to Modified Causal Forest

Modified Causal Forest is an extension of the original Causal Forest framework by introducing some modifications within the standard causal forest methodology. These modifications are aimed at improving the accuracy of treatment effect estimation.

- **First modification** is the improvement in sample splitting by introducing new splitting rule of trees that can reduce the selection bias in observational studies.

- **Second modification** exploits a weight-based inference procedure that allows for flexible aggregation of treatment effects across different levels (i.e. individual (IATEs), group (GATEs), population (ATEs)) making it easier to interpret and apply results in practice.

Advantage of MCF over Traditional Approach

- Traditional approach typically relies on semi-parametric methods whereas machine learning approach such as MCF are non-parametric. Therefore, it is more flexible.

- As traditional matching methods estimates subgroup effects (GATEs) independently, **it's not possible to jointly assess and compare GATEs alongside the overall Average Treatment Effect (ATE).** Therefore, it is not feasible to detect treatment effect heterogeneity using matching methods.

  - In contrast, **MCF overcomes this challenge by enabling the joint estimation of GATEs.** MCF can model treatment effect variation across subgroups, allowing for a more robust and data-driven analysis of effect heterogeneity.

Implementing MCF Framework

Case Study:

- A country runs a training program for unemployed individuals to help facilitate their transition back into the labour market. It is intended to support them finding a job and attain higher earnings profiles by raising their skill level, particularly of the low skilled. The government is interested in estimating the effects of the program on the earnings of unemployed individuals following their participation. Here are some specifications of the program:
  - **Duration**: 3-6 months.
  - **Start of the program**: 1st quarter, 1993
  - **Age group**: 30-50 years
  - **Data**: Rich pre-participation information and post-participation outcome.
- **Comparison group** consists of individuals who are eligible for participation but did not participate and have similar socio-demographic characteristics.

Steps for MCF Implementation

Step 1: Institutional setting and context of evaluation

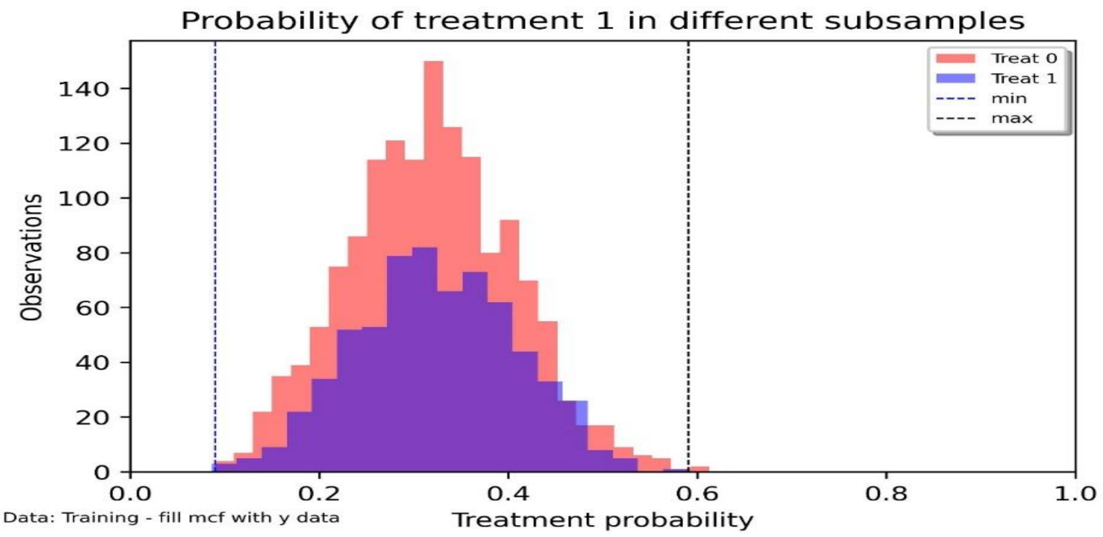
- This is an observation study on labour market program. Rich administrative data is available on short duration Skills Development training program.
- The main objective of this impact analysis is to estimate the average treatment effect of the treated and heterogenous treatment effects of the training program on medium term outcome using causal machine learning MCF approach.

Step 2: Causal Modelling

- Treatment: labour market training program
- Outcome: Earnings during the 4th quarter of 6th year post participation
- Sufficiently large number of socio demographic and labour market information. The model is based on selection on observables.

Step 3: Identification Assumptions

- **CIA**: It holds as there is a large number of socio demographic characteristics and labour market history.
- **Common Support**: We can observe very good overlap between treatment group and comparison group in raw data.
- **SUTVA**: No spill over effect as the program is small compared to the relevant regional labour market.
- **Exogeneity of Confounders:** Treatment unlikely to have any influence on the confounding variables.

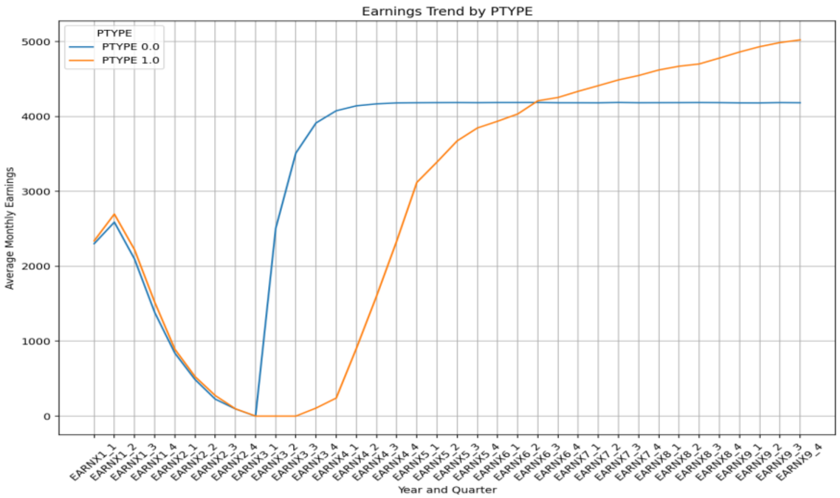


Step 4: Data Processing and Comparison Group Selection

- **Data processing*:** Removing missing values, checking for duplicates, as well as variable transformation. to prepare the final analytical file.
- **Comparison group Selection:** Comprise of eligible non-participants. Ensure they have the same information as participants. Need to have sufficiently large comparison group.

Step 5: Pre-estimation Diagnosis

Outcome Trend Analysis:
We track the earnings of participants and comparison group before and after program participation. For this short-term training program, we observe a lock in effect during the first two quarters of participation.



*Data consists of 13628 individuals with 4251 participants and 9377 comparison group. Total number variable is 82.

Step 6: Estimating IATE and Average Treatment Effect

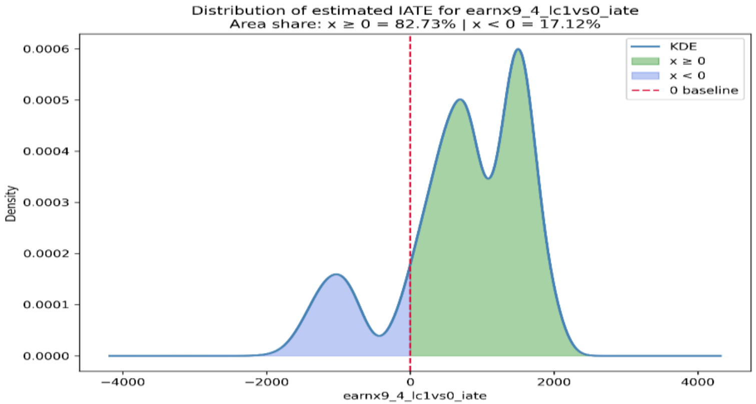
- MCF first generates the most granular level of treatment effect estimate referred to as IATE.
- The average impact of the program on the population (ATE) and participant (ATT*) can be obtained by aggerating IATEs. The estimated values of ATE and ATT for the case study are highly statistically significant.

Parameter	Estimate	P value
ATE	723	0
ATT	662	0

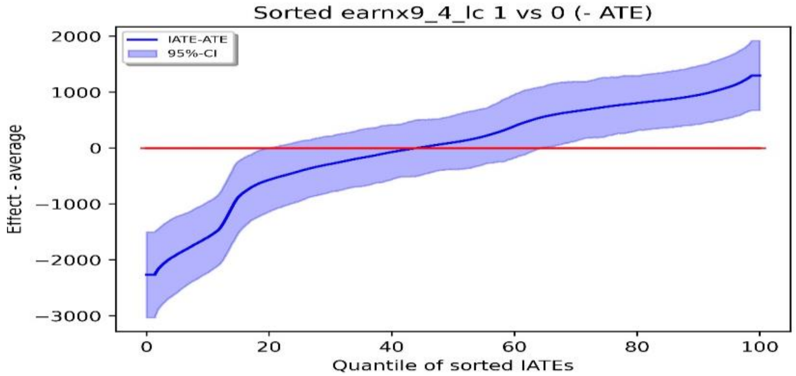
*In this study, we mainly focus on the estimates that are related to the treated population (such as ATT, GATET, BGATET) as we are mainly concern of the impact of a treatment or intervention on those who actually received it.

Step 7: Identifying Heterogeneity at Individual Level

- **Distribution of the Estimated IATEs**: Analyzing the density plot of the estimated IATEs, we observe that most (83%) of participants are **better off.** However, there are some participants (17%) who are **worse off** as they have negative IATEs relative to similar non-participants.



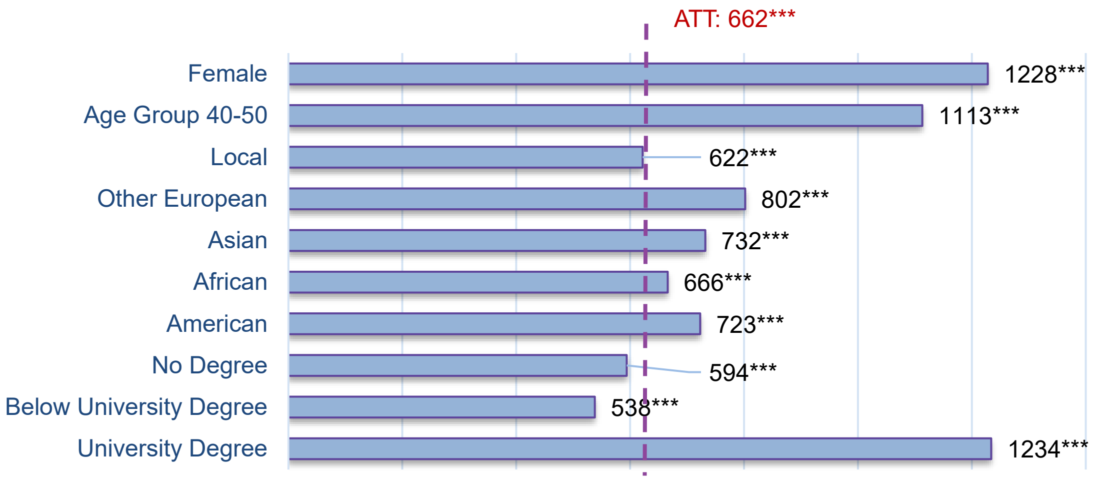
- **Plotting IATE-ATE:** To detect effect heterogeneity, we plot the difference between IATEs and ATE. We can observe that there is gap between IATE-ATE line and zero reference line (as well as no overlap between the confidence intervals). This provides evidence that there is heterogeneity in treatment effect at the individual level.



Step 8: Identifying Heterogeneity at Group Level

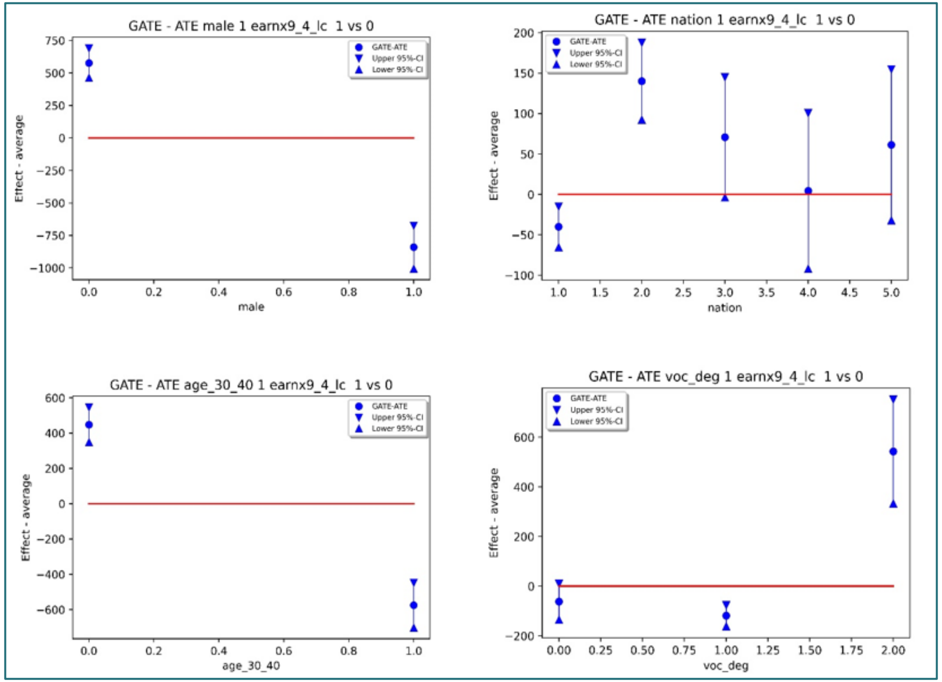
To identify the effect heterogeneity, we first estimate the **Group Average Treatment Effect for Treated (GATETs)** for four socio demographic characteristics: **gender, age, nationality, education level**.

All subgroups have statistically significant GATET except male and age group 30-40.



Step 8 (cont'd): Detecting Heterogeneity (GATET-ATT)

- To detect effect heterogeneity, we need to check if the difference between GATETs and ATT is statistically significant.
- There is evidence of effect heterogeneity for some of the subgroups as the GATET-ATT estimates are statistically significant for them.



Step 9: Interpreting Treatment Effect Heterogeneity

- Once we detect effect heterogeneity, the next step is to identify the source of heterogeneity.

- In other words, we are interested to better understand if this effect heterogeneity is arising due to the **variable of interest** or the **confounding features**.

- To illustrate how to interpret the effect heterogeneity, we focus on **European nationality** and **female participants** subgroups.

- There are two approaches to interpret this heterogeneity
  - Analyzing Profile of Subgroups
  - Balancing Group Average Treatment Effects on Treated (GATETs)

Step 9 (cont'd): Balanced Group Average Treatment Effects (BGATEs)

- When we detect effect heterogeneity using GATET-ATT estimate, the covariates across subgroups (based on a feature such as gender) are not balanced.

- If we balance covariates of the subgroups while estimating GATETs, we get a new estimate known as **Balanced Group Average Treatment Effects of the Treated (BGATETs).** To detect effect heterogeneity using BGATET-ATT estimate, we can follow the rule of thumb below:

- **Rule of Thumb:**
  - Once we balance the GATET with respect to all the covariates* and it turns out that the estimate of BGATET-ATT becomes **statistically insignificant**, we can conclude that there is **no evidence of effect heterogeneity.**
  - However, if the BGATE-ATT remains **statistically significant** even after balancing all the covariates, then we can conclude that there is **evidence of effect heterogeneity** for that variable of interest.

* The confounders can be correlated, so it is difficult to pin-point the covariates that are contributing to the effect heterogeneity the most. As our policy interest is to identify whether the main source of effect heterogeneity is coming from the variable of interest or from the confounding factors, we suggest to balance all the covariates and then check if the effect heterogeneity still exist.

Step 9 (cont'd): Interpreting Effect Heterogeneity

- **Non- European:** As the BGATET-ATT is not statistically significant after balancing all the covariates, so we can conclude that there is **no evidence of effect heterogeneity** based on nationality for the non-European subgroup.

- **Female:** As the BGATET-ATT is statistically significant after balancing all the covariates, so we can conclude that there is some **evidence of effect heterogeneity** based on gender for the female subgroup

Parameter	Balancing Variable	Non- European		Female	
		Co-efficient Estimate	P-value	Co-efficient Estimate	P-value
GATET-ATT	None	140	0	566	0
BGATET-ATT	All covariates (Labour Market Information plus Socio-Demographic Characteristics)	63	**0.15**	700	**0**

Step 10: Post estimation analysis using clustering

- To understand **which group benefits the most** we conducted K means clustering. The findings suggest we have **3 clusters**.
  - The participants in the **third** cluster benefits the most.
  - The cluster/group that benefits the most can be characterized as mainly **female, aged between 40-50, university graduate** and **having high labour market attachment prior participation.**

VARIABLES	CLUSTER 1	CLUSTER 2	CLUSTER 3
ESTIMATED IATE	-1034.85	542.64	**1511.85**
NUMBER OF OBSERVATIONS	792	2382	**2264**
AGE_30_40	0.98	0.53	**0.16**
MALE	1	0.54	**0.04**
EDUCATION LEVEL	0.85	0.96	**1.22**
EARNX1_1	2086.74	2239.39	**2468.69**
EARNX1_2	2256.3	2509.83	**2840.78**
EARNX1_3	1899.65	1995.95	**2285.88**
EARNX1_4	1267.68	1310.99	**1573.79**

Conclusion

- The MCF algorithm is an alternative approach to traditional matching methods for estimating labour market program impact.

- There are several advantages of MCF over matching approach especially the detection of effect heterogeneity which is not feasible under matching methods. Detecting and understanding effect heterogeneity can have important implications for policy development.

- Implications of this study can also be extended to other fields such as **social programs, business decisions, marketing campaigns, or medical treatments** when the objective is to measure treatment effect heterogeneity of an intervention.

Thank You!

Questions?

jamil.sayeed@hrsdc-rhdcc.gc.ca
andy.handouyahia@hrsdc-rhdcc.gc.ca
essolaba.aouli@hrsdc-rhdcc.gc.ca