Digital Inclusion: Understanding e-Vulnerability in Canada

Firouz Fallahi

Service Policy and Strategy Directorate
Employment and Social Development Canada (ESDC)

Quantitative Impact Assessment Workshop, Feb 2026

Outline

- Background
- Definition of the e-vulnerability Index (EVI)
- Data Sources and the Associated Challenge
- Methodology
  - Principal component analysis (PCA)
  - Small Area Estimation (SAE)
- Modeling and Calculation of EVI's Subcomponents
- Results
- Summary
- Policy Implications

Context – Digital Transformation of Service Delivery and Disparities in Access to Services

- The Government of Canada is increasingly enhancing its online delivery of services and information

- Our digital government transformation aims to modernize and adapt service delivery to better meet the needs of Canadians in every part of the country

- The increasing focus on digital communication and interaction underscores the importance of the need for government to proactively identify citizens who may be more likely to face barriers to accessing and using online services. We refer to this group as the ***e-vulnerable***

Rationale of the Project

- e-vulnerable people are at greater risk of being at a service disadvantage, and thus are more likely to experience more difficulties to obtain the service they need online
- Additionally, digital divide exacerbates other socio-economic disparities, further widening these disparities; therefore, closing digital gaps is key to achieving the UN's Sustainable Development Goals as well

The aim of this study is to identify e-vulnerable population subgroups and to determine their geographical locations

- Which Canadians are most likely to be e-vulnerable?
- How does e-vulnerability vary across different sub-groups of the population and across different geographic areas?

What Is the E-vulnerability Index (EVI)?

- Employment and Social Development Canada (ESDC) has developed its e-vulnerability Index (EVI) to measure digital vulnerability for the first time in 2015 and has undergone multiple updates since

- The EVI examines three subdimensions of e-vulnerability:

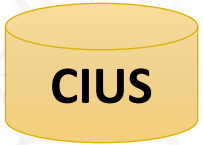
1. ***Access:*** examines whether individuals have the necessary means to benefit from the Internet
2. ***Competencies:*** captures whether individuals have the necessary skills and knowledge to take full advantage of the Internet
3. ***Comfortability:*** considers the willingness or desire of individuals to use the Internet

Creating the EVI: Data Source

Data on the Internet use of the Canadian population is primarily collected from Statistics Canada datasets:

➢ Canadian Internet Use Survey (CIUS)

- Most recently, the survey was conducted in 2022 across 10 Canadian provinces
- Sample size: about 25 000 individuals aged 15 and older
- Includes a set of questions on:
  - Internet access and use
  - Activities performed online, from sending and receiving emails to using governments services online
  - Difficulties, security issues, or privacy concerns related to using Internet



- The CIUS was used to measure the three subdimensions of the EVI, i.e., **Access**, **Comfortability**, and **Competencies***

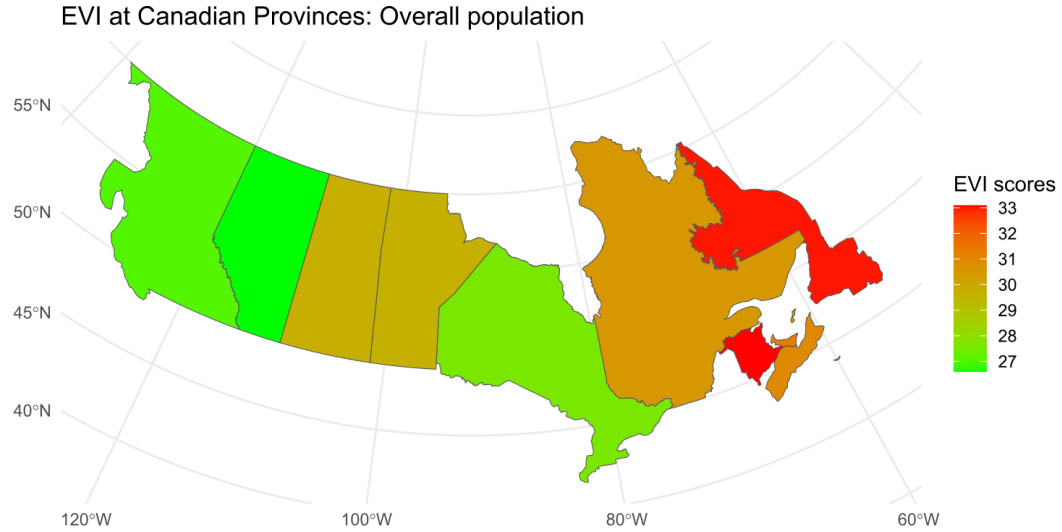
* In earlier versions of the EVI, competencies were evaluated based on data from the Programme for the International Assessment of Adult Competencies (PIAAC). However, due to the removal of the Problem Solving in Technology-Rich Environments module in the latest PIAAC cycle, data from the CIUS was used instead.

Creating the EVI: Methodology (I)

Principal Component Analysis (PCA)

- The CIUS includes several indicators that are relevant to one or more of the EVI dimensions (access, comfort, or competency)
- Principal Component Analysis is used to reduce the full set of Internet-related indicators present in the CIUS to a manageable subset and to measure the 3 subdimensions of the EVI
- Using this approach, we obtain reliable EVI scores, but only for limited set of geographies as shown on the next slide

Creating the EVI: The Challenge



- The CIUS sample size is sufficient to produce reliable estimates at the national and provincial level; however, the precision decreases when analyzing smaller communities or demographics

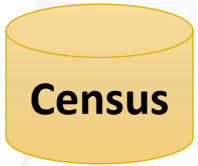
Creating the EVI: The Challenge (cont'd)

Challenge:

Lower precision at smaller or finer levels of aggregation makes it ineffective for evidence-based policymaking and for targeted interventions

Solution:

Census of Population



- The 2021 Census of Population gathers extensive demographic and socio-economic data for millions of individuals and allows to do estimations at very small area levels and for various sociodemographic groups
- **But it lacks data on Internet use, access, and skills**

Creating the EVI: Methodology (II)

- To address this gap, we utilized **Small Area Estimation Techniques** to estimate the three subdimensions of e-vulnerability for everyone in the Census
- This approach produces accurate, granular, and highly disaggregated information, better supporting policymaking, targeted interventions, and other public policy applications
- Each Census respondent is assigned an EVI score, enabling tailored results for various geographic regions and demographic segments, e.g., seniors and immigrants

What is Small Area Estimation (SAE)?

- SAE is concerned with the development of statistical procedures for producing precise estimates for small areas, i.e. for domains with small or zero sample sizes (Pratesi and Giusti, 2015)
- Rao and Molina 2015 define small area as "any area/domain for which direct estimates of adequate precision cannot be produced“
  - It combines data from 2 or more sources to provide reliable estimates
  - It is widely used in official statistics when sample sizes are too small to provide reliable estimates at detailed levels
  - It **borrows strength** from related areas and from auxiliary dataset(s)
  - It provide indirect estimates when direct estimates are not feasible

SAE Approaches

There are 2 main approaches used in SAE:

1) Area level models:

- Using direct estimation between variable of interest and auxiliary variables at area level
- The most popular model is Fay and Herriot's (1979)

$$y_i^{Dir} = x_i^T \beta + u_i + e_i$$

  - Where y_i is the parameter of interest at domain i and Dir shows direct estimates
  - e_i is the sampling error
  - u_i is the random effect and assumed to be independent of sampling error

SAE Approaches (cont'd)

- This approach has a few pros and cons, some of them are as follows:
  - Pros: It requires area-level data rather than individual-level data, which is more widely available and not subject to strict confidentiality or privacy rules
  - Cons: Area-level results are generally less efficient, and they cannot be broken down into smaller units once generated

SAE Approaches (cont'd)

2) Unit-level models

- In this approach, unit-level data (e.g. individual or household) is used, which is much richer than area level data but requires access to microdata
- We use unit-level models to estimate the EVI subdimensions:

$$y_{i,j} = x_{i,j}^T \cdot \beta + \delta_i + e_{ij}$$

$y_{i,j}$ shows the EVI subdimension (Access, Comfortability, or Competency)

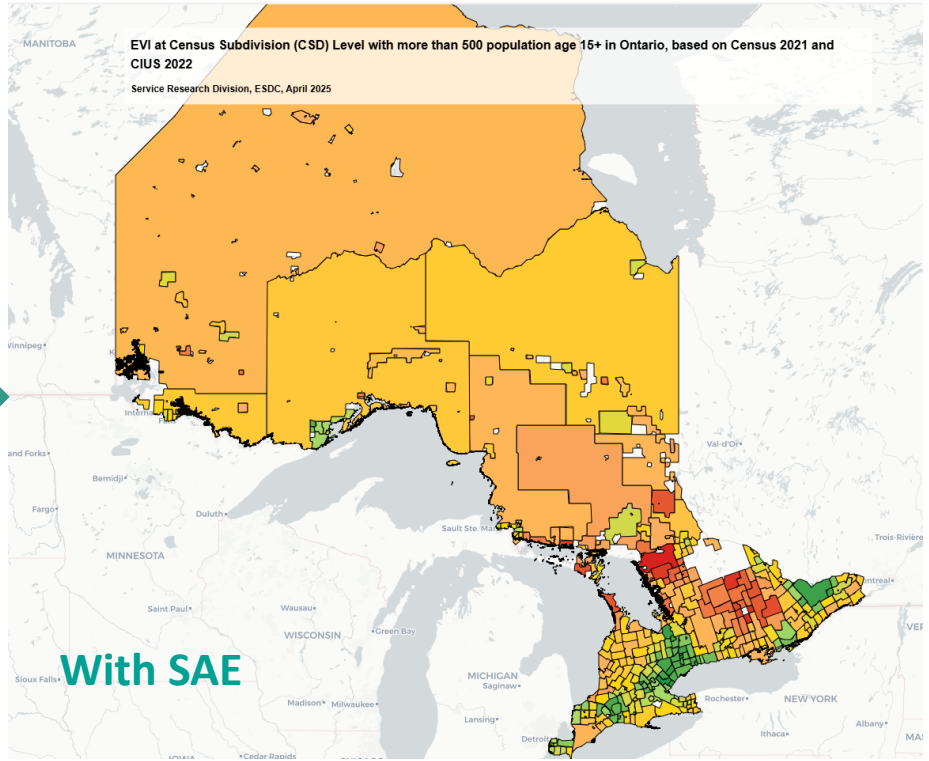
$x_{i,j}^T$ represents a vector of explanatory auxiliary covariates for respondent j in location i,

e_{ij} are assumed to be independently distributed and are independent of δ_i

δ_i is the area specific random-effect to capture the difference between areas that is not explained by $x_{i,j}^T$.

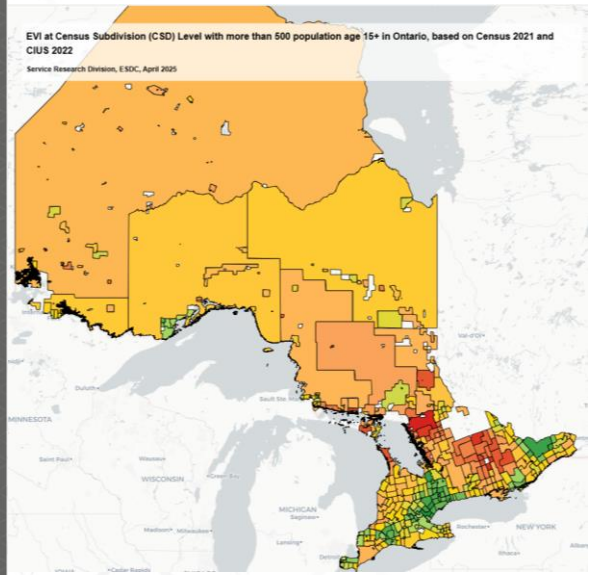
Comparison of Estimates: No SAE vs. SAE

Without SAE

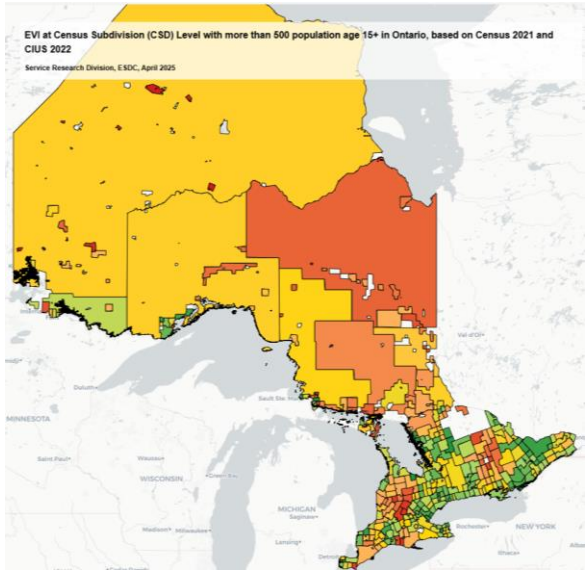


With SAE

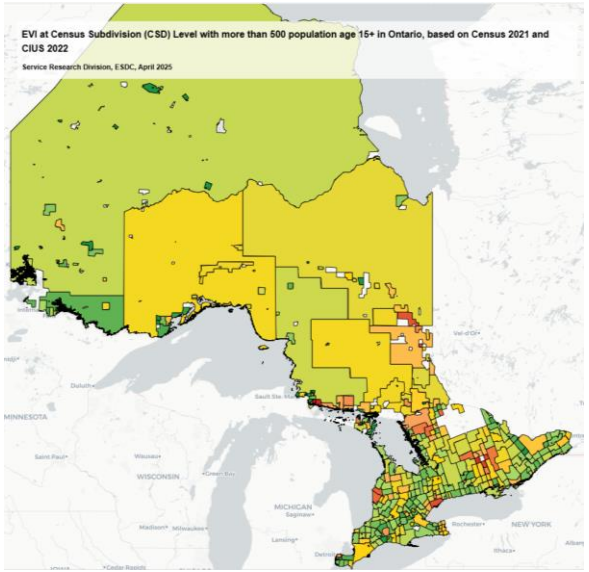
SAE: Building Insights through SAE Capabilities



EVI scores

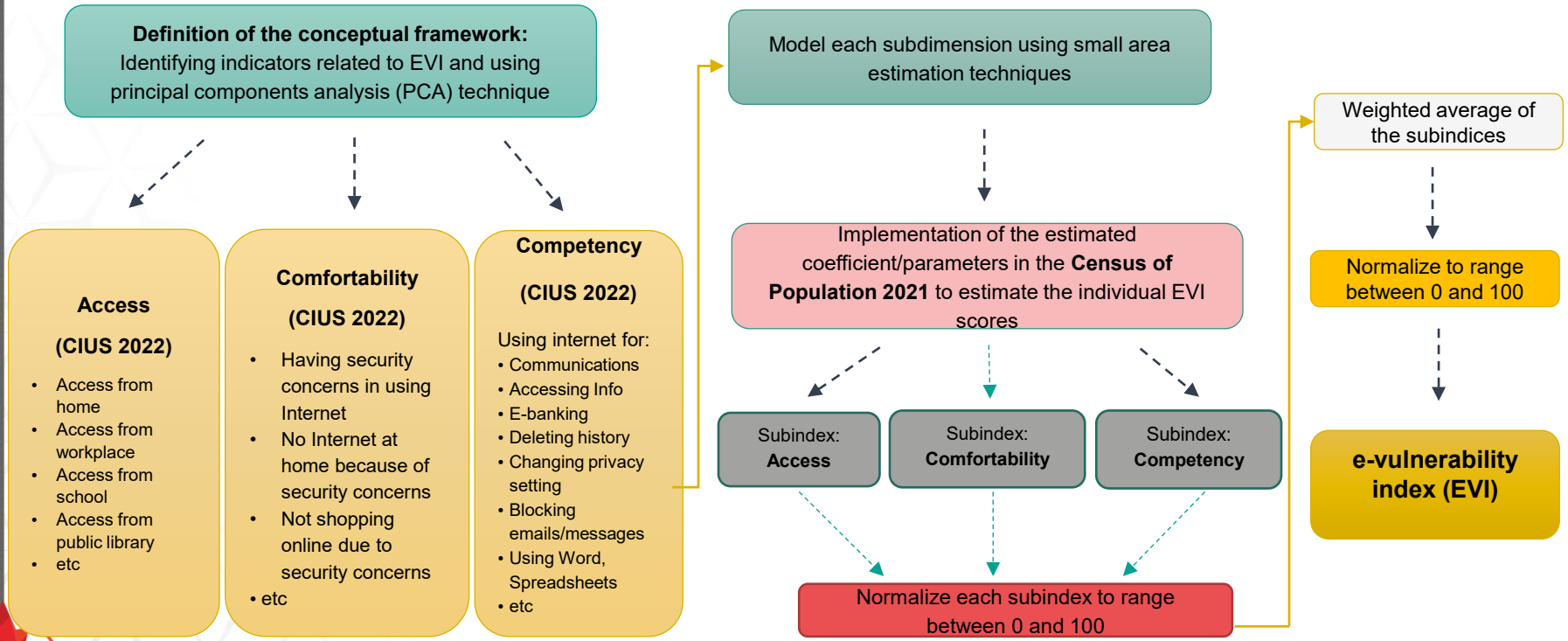


Less than high school diploma, Aged 15 and over



In low-income (Based on Market Basket Measure (MBM))

An Overview of the Creation of the EVI

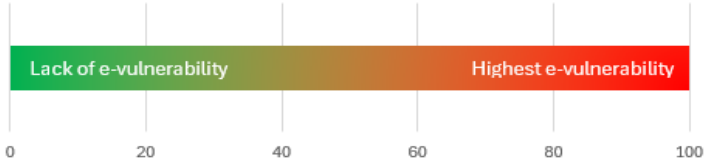


Estimating the Subdimensions of the EVI (I)

- The **number of principal components** in the CIUS was established by analyzing eigenvalues and the cumulative variance explained
  - The weights assigned to each component were based on the percentage of total variance explained by each component
- **Unit-level models** were developed using a Bayesian Markov Chain Monte Carlo (MCMC) approach
  - This included thinning and burn-in procedures
  - Convergence and lack of autocorrelation in the MCMC samples were confirmed
- **Mapped into Census** and applied the coefficients estimated from the MCMC analysis to calculate the access, comfortability, and competency scores for **each individual** in the 2021 Census

Estimating the Subdimensions of the EVI (II)

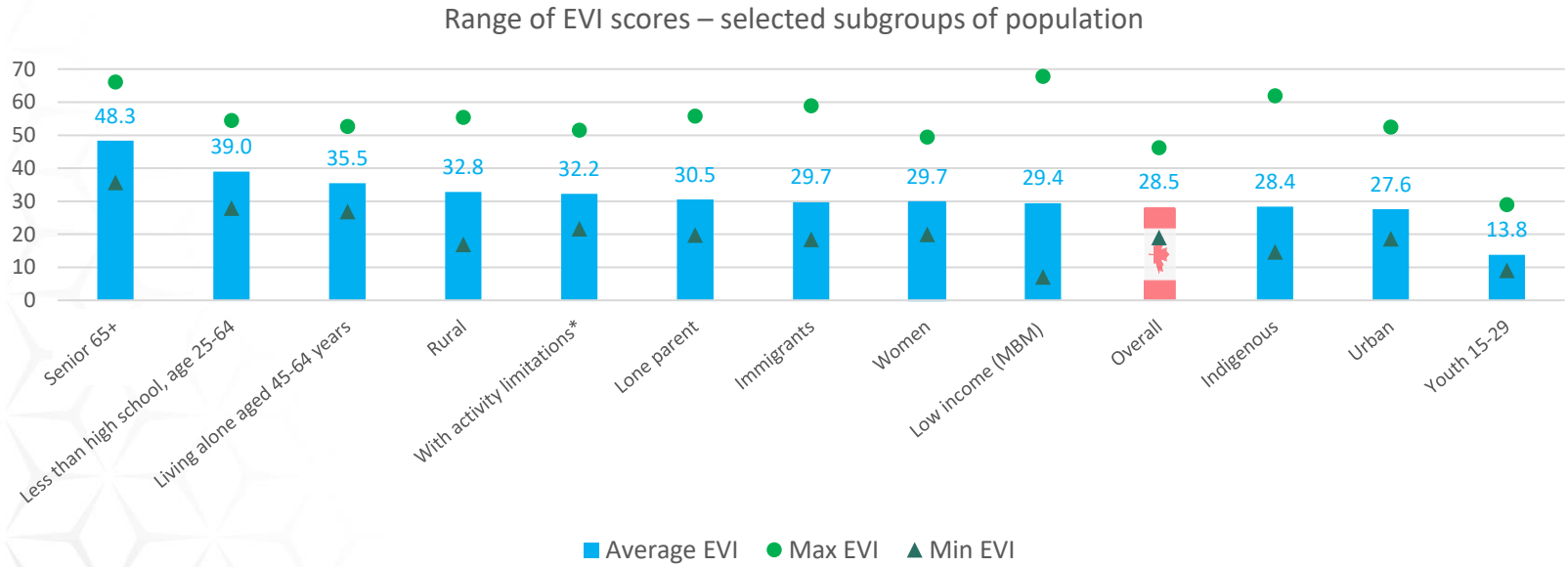
- The e-vulnerability **subindices were normalized** to rescale them to be between 0 and 100
- **Equal weights** were assigned to each subdimension of the e-vulnerability and aggregated to get the EVI score for each individual
  - Sensitivity analyses were conducted by varying the weights; however, the overall conclusions stayed broadly the same
- The results for the **EVI scores were normalized** again to have the final EVI scores ranging between 0 and 100



- At the end, EVI scores were available for over 7 million (unweighted) respondents in the Census
  - These scores were estimated at individual level which allows to aggregate at finer levels of geography for subgroups of population in Canada

E-Vulnerability Index: Results

- E-vulnerability varies based on different dimensions of a person's life. For example, age, immigrant status, geographic location and activity limitations can all effect ones' degree of e-vulnerability

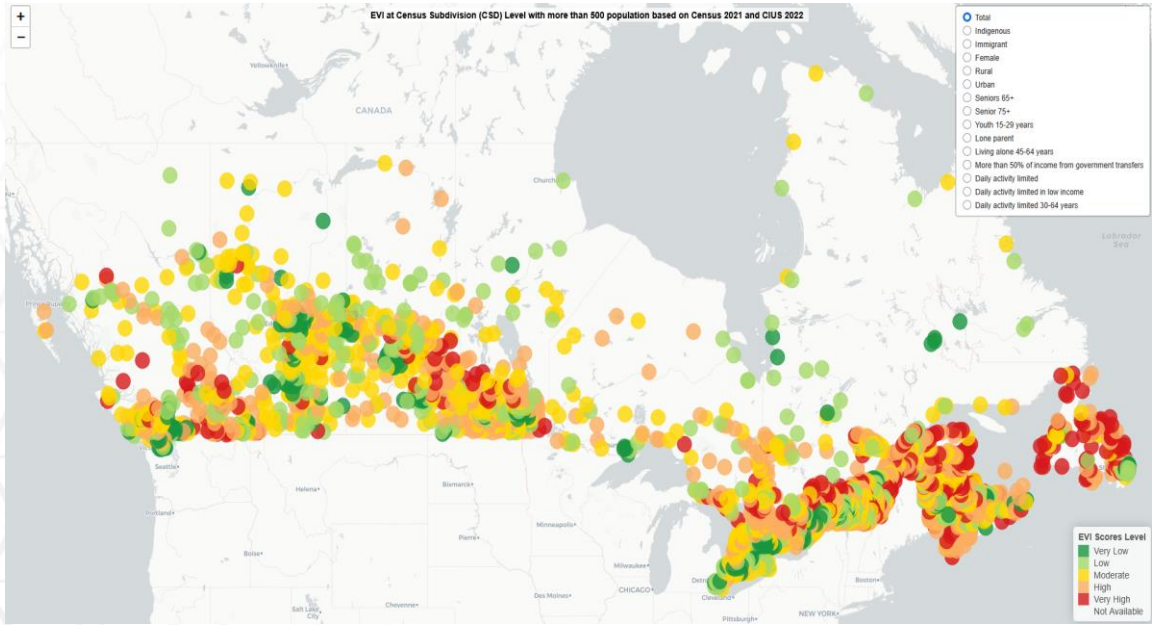


** As defined in the Census of Population. People with activity limitations represent a broader group than people with disabilities as defined by ESDC based on the Canadian Survey on Disability.*

E-Vulnerability Index: Results

- In Canada, the overall EVI score is 28.5, serving as a benchmark for comparing the e-vulnerability of population subgroups against the average
- Among the subgroups that are particularly relevant to ESDC:
  - **Seniors**, **individuals with activity limitations**, and those residing in **rural** areas have a higher level of e-vulnerability compared to the average
  - Conversely, **youth** have the lowest EVI score
  - There is notable variation in EVI scores within each subgroup:
    - for seniors, average scores can range from 36 to 66, indicating that they are generally more e-vulnerable regardless of their location or socio-economic status
- EVI scores are location-dependent, as will be shown on the next slides

Overall population at Census Subdivision (CSD) level

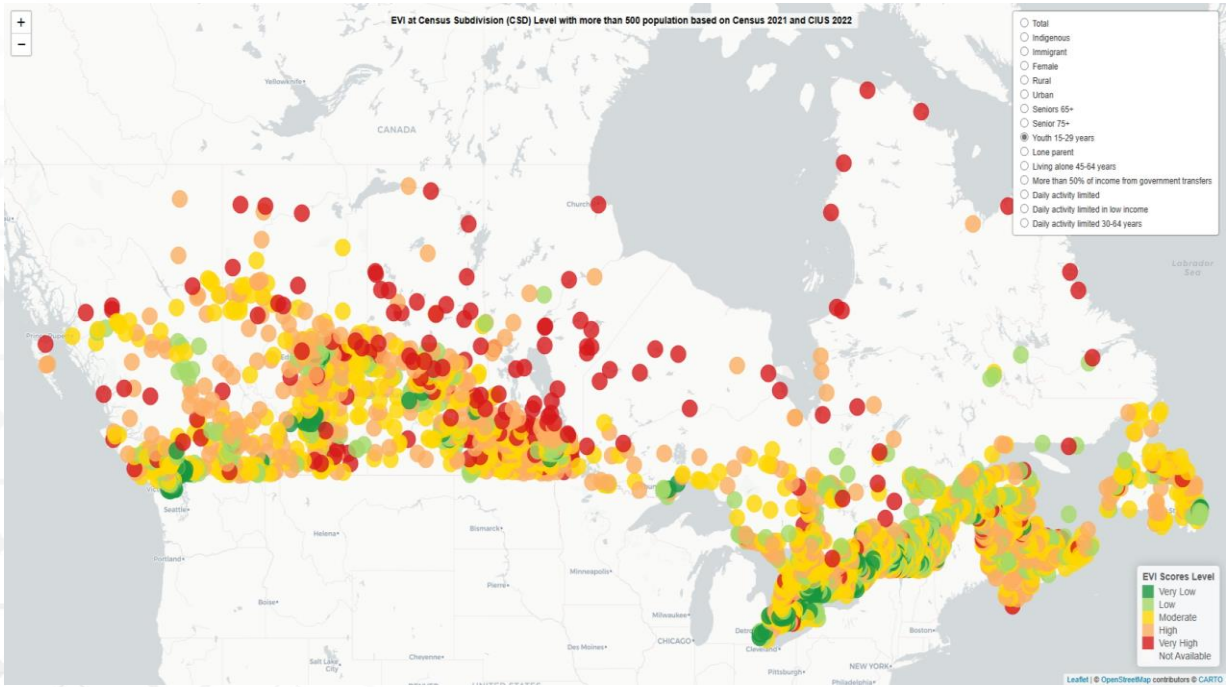


The distribution of EVI scores across geographic areas vary significantly, depending on which sub-group we look at

For the overall population, most of the CSDs with a high EVI score (red dots) are located in the Eastern provinces

However, this does not hold for all population subgroups

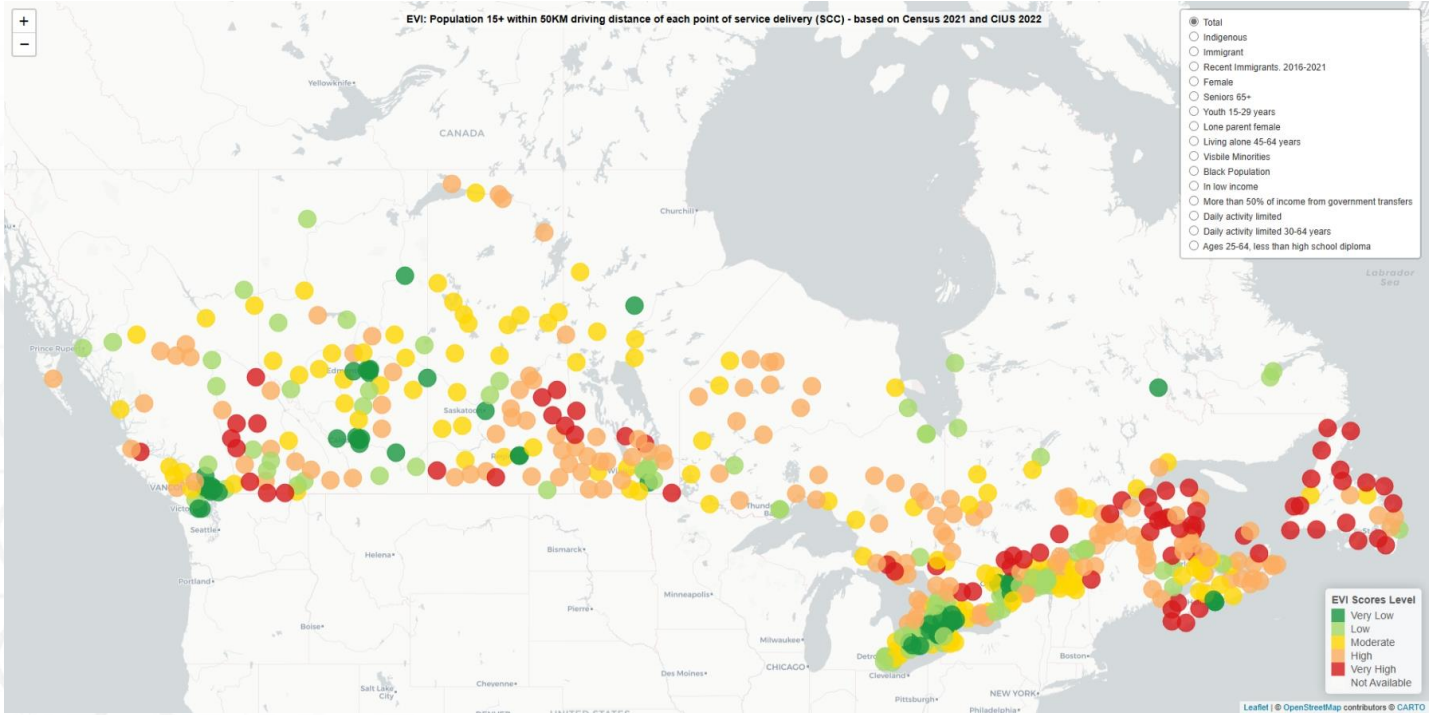
Youth 15-29 years of age at CSD level



- Youth have the lowest EVI scores (13.8) among demographic groups
- Their EVI scores range from 9 to 29
- The highest EVI scores are mainly in Manitoba and Saskatchewan, particularly in their northern areas

Service Canada Points of Service Delivery (SCC):

50KM driving distance



The EVI score for any specific SCC is driven by the characteristics of its population (the next slide presents this info for the SCCs with the highest and lowest EVI scores)

CSDs where SCCs with the highest and lowest EVI scores are located: Selected Characteristics

CSD: Springdale (NL) Population 2,965 Average EVI score of population 40.2	**Socio-demographics characteristics of the population in the CSD:** ▪ High proportions of seniors in population (31.7%) ▪ High low-income (LIMA) prevalence rate (21.5%) ▪ Lower employment rate (37.3%) ▪ Higher unemployment rate (15%) ▪ Higher proportion of population with less than High School Diploma (33.5%) ▪ Living alone (12.3%)
CSD: Canmore (AB) Population 15,990 Average EVI score of population 24.0	**Socio-demographics characteristics of the population in the CSD:** ▪ Low proportions of seniors in population (17.4%) ▪ Lower low-income (LIMA) prevalence rate (6.5%) ▪ Higher employment rate (65%) ▪ Lower unemployment rate (9.7%) ▪ Lower proportion of population with less than High School Diploma (6.8%) ▪ Living Alone (11.6%)

Summary

- The ongoing shifts in socioeconomic and technological environments highlight the need to improve data collection methods to better reflect our changing data requirements
- Gathering detailed data is costly, so it is more effective to integrate data from multiple aspects of individuals' lives and activities
- We employed the Small Area Estimation (SAE) technique, which helped address the CIUS sample-size limitations in a cost-effective way by combining data from the 2021 Census of Population
    - It enabled us to compute EVI scores at the individual level, allowing for tailored results that can be aggregated according to specific geographic regions and demographic groups

Policy Implications of the EVI

- The EVI helps to identify the specific areas and demographics where alternative methods of service delivery (such as mail, phone, or in-person) or information dissemination (including radio or newspapers) are especially required
- Facilitate outreach by informing outreach officers of an area's e-vulnerability before conducting in-person visits
- Aids in developing digital services in such a way that mitigates and addresses potential drawbacks that may result from the introduction of e-services
- A uniform one-size-fits-all policy will not be effective

Thank You!

Service Research Division, Service Policy and Strategy Directorate, ESDC

NC-DPSS-Recherche_Services-Service_Research-SPSD@hrsdc-rhdcc.gc.ca