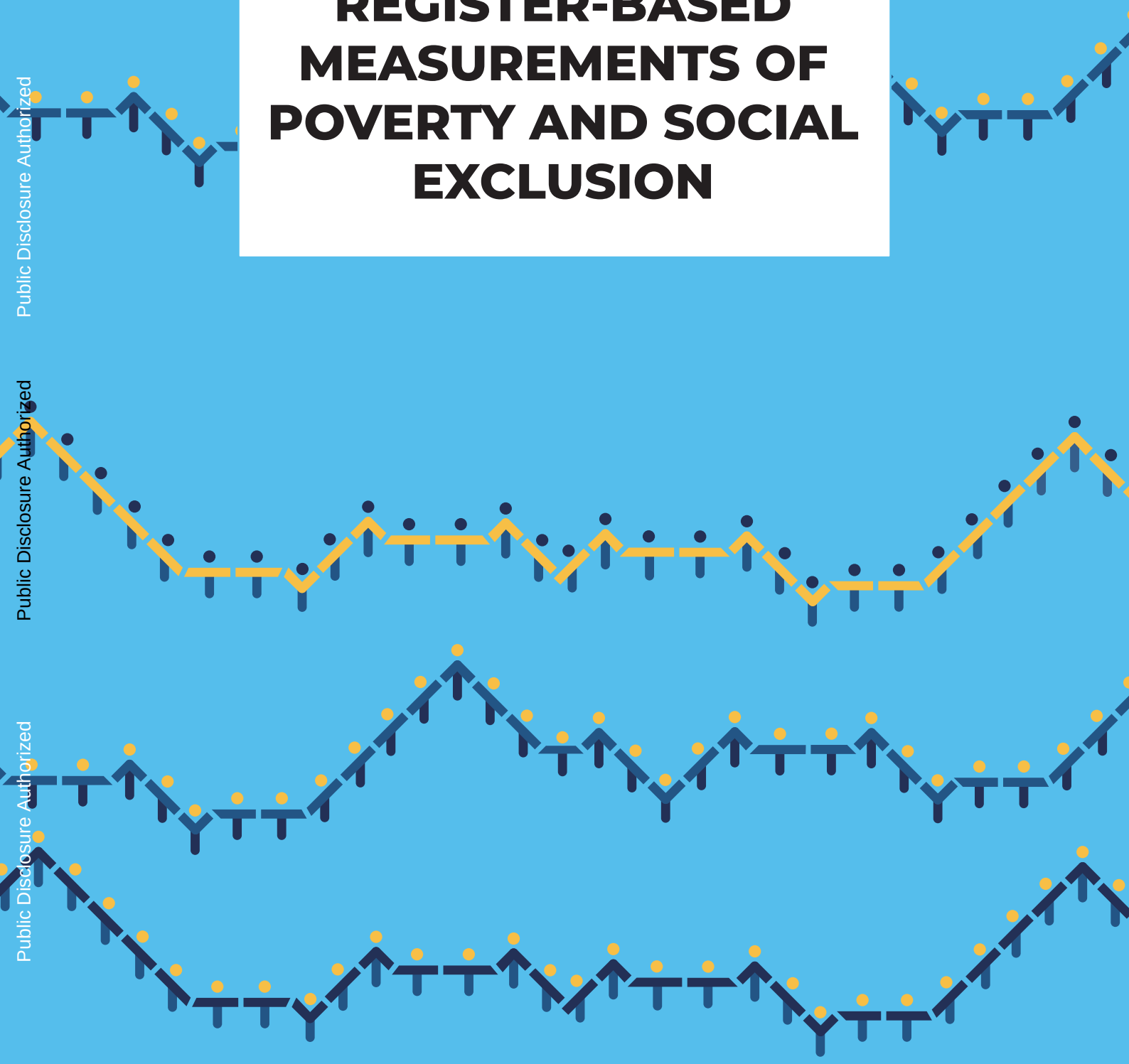
# REGISTER-BASED MEASUREMENTS OF POVERTY AND SOCIAL EXCLUSION



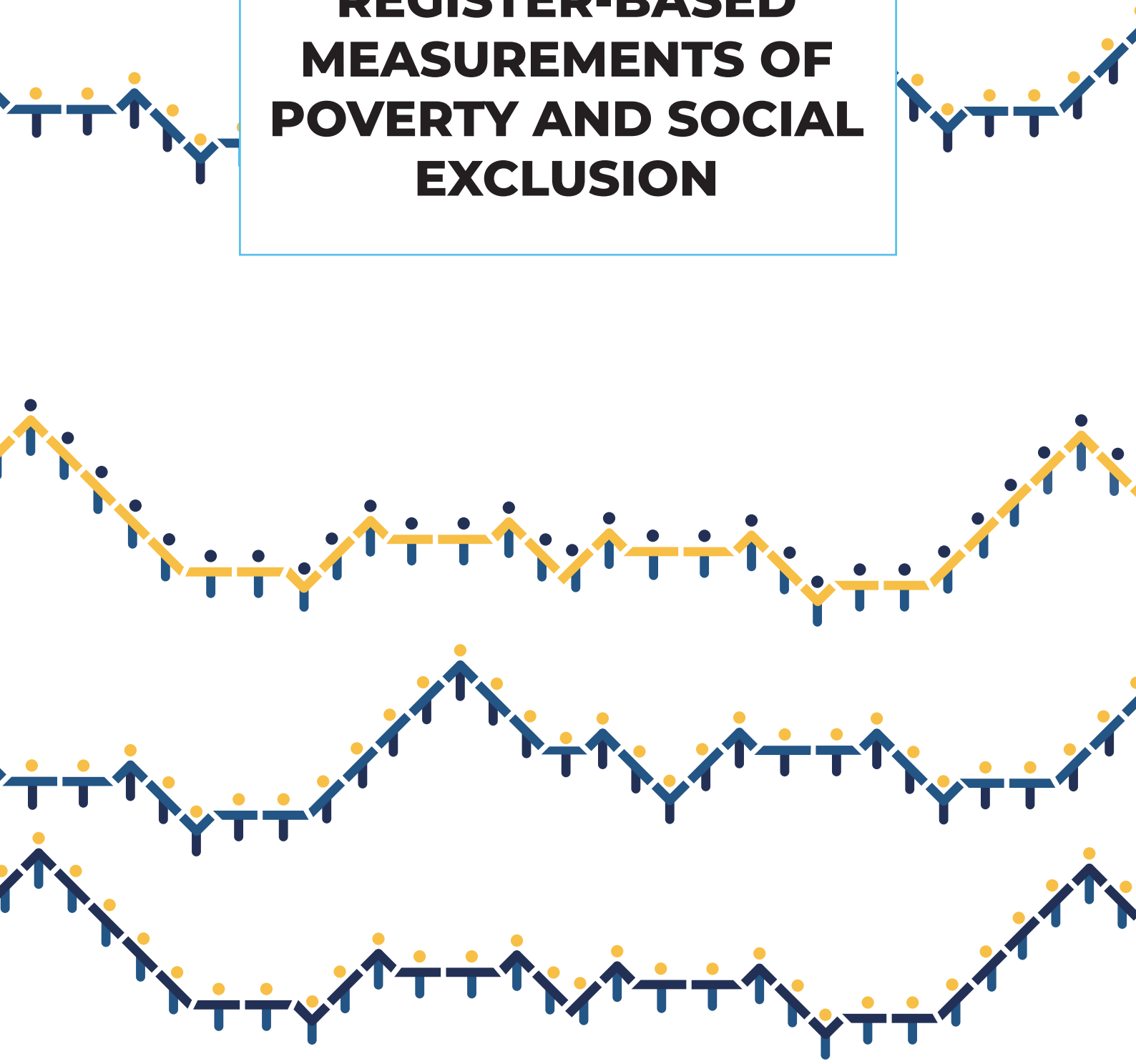
Funded by the European Union

REPUBLIC OF CROATIA
Ministry of Labour, Pension System, Family and Social Policy

THE WORLD BANK
IBRD • IDA | WORLD BANK GROUP

Croatia Technical Support Instrument
Support to Improve Poverty Diagnostic and Monitoring of Social Policies

# REGISTER-BASED MEASUREMENTS OF POVERTY AND SOCIAL EXCLUSION



June 2024

# DISCLAIMER

This document was produced with the financial assistance of the European Union. The views expressed herein can in no way be taken to reflect the official opinion of the European Union.

This report is a product of the International Bank for Reconstruction and Development / The World Bank. The findings, interpretation and conclusions expressed in this paper do not necessarily reflect the views of the Executive Directors of the World Bank, the European Commission, or the Government. The World Bank does not guarantee the accuracy of the data included in this work.

# COPYRIGHT STATEMENT



# RIGHTS AND PERMISSIONS



This report has been delivered in June 2024, under the EC Contract No. REFORM/IM2022/004 under TF073805, signed between the European Commission and the International Bank for Reconstruction and Development. It corresponds to Output 1.2. for the Technical Assistance on Support to Improve Poverty Diagnostic and Monitoring of Social Policies in Croatia under the above-mentioned agreement.

# CONTENT

## ANNEXES 94



## REFERENCES 114

## LIST OF BOXES

## LIST OF FIGURES



## LIST OF TABLES

# ACKNOWLEDGEMENTS

This report was prepared by the World Bank (WB) in close cooperation with the Croatian Ministry of Labor, Pension System, Family, and Social Policy (MLPSFSP) and the European Commission's Directorate General for Structural Reform Support (DG REFORM). The report is delivered under *Croatia: Support to Improve Poverty Diagnostic and Social Policies Monitoring* task financed by the European Union via the Technical Support Instrument.

The World Bank team preparing this report consisted of Nga Thi Viet Nguyen (Task Team Leader, WB), Ana Šimundža (Operations Officer, WB), Martina Vojković (Consultant, WB), Ivica Rubil (Consultant, WB), Marko Krištof (Consultant, WB), Vedran Recher (Consultant, WB), and Nataša Filipović (Consultant, WB). Administrative support was provided by Bogdanka Krtinić (Program Assistant, WB), while Marcello Arrigo edited the report and Anamarija Perić delivered the graphic design. The team benefited from a rich series of comments from peer reviewers Andrew Dabalen (Chief Economist, WB), Zurab Sajaia (Senior Economist, WB), and Roy Van der Weide (Senior Economist, WB). The team appreciated additional materials and insights from Reena Badiani-Magnusson, Samuel Freije-Rodriguez, Vlad Alexandru Grigoras, Dean Mitchell Jolliffe, Daniel Gerszon Mahler, Paul Scott Prettitore, and Anja Sautmann (all from WB). Special thanks were given to Rema Hanna (Harvard University) for her review of this report.

The team was grateful for invaluable comments from Milena Koren (Project Team Leader for Components 1 and 2 on behalf of the MLPSFSP), Rima Joujou Deljkić (DG REFORM), and representatives from the Croatian Bureau of Statistics, the Tax Administration of the Ministry of Finance, and the Ministry of Justice, Public Administration, and Digital Transformation. The team appreciated insights from Katica Lažeta, Nikola Vinski, Sandra Batlak, Monika Begović, Marija Kalečak, Branka Zečević, Željka Grbac and Gordana Kučan Šarfi (all from the MLPSFSP). This report was enriched by feedback from participants of the consultative workshops held on January 29, 2024, and July 1, 2024, in Zagreb, and by the experiences shared by the governments of Estonia and Slovenia during study visits in February and June 2024.

The team worked under the guidance of Marina Wes (WB Country Director for Bulgaria, Croatia, Poland, and Romania), Jehan Arulpragasam (WB Country Manager for Croatia), and Ambar Narayan (WB Poverty and Equity Practice Manager).

# ABBREVIATIONS

| | |
|---|---|
| **APIS-IT** | Information System and Information Technologies Support Agency |
| **AROP** | At risk of poverty |
| **AROPE** | At risk of poverty or social exclusion |
| **CATI** | Computer-assisted telephone interview |
| **CBS** | Croatian Bureau of Statistics |
| **CES** | Croatian Employment Service |
| **CPI** | Consumer Price Index |
| **DG REFORM** | European Commission's Directorate-General for Structural Reform Support |
| **DI** | Development Index |
| **EC** | European Commission |
| **EDIP** | Electronic Records of Income and Receipts |
| **ENC** | Enabling Condition |
| **ENDIREH** | Mexico National Survey on the Dynamics of Household Relationships |
| **EPSR** | European Pillar of Social Rights |
| **ERDF** | European Regional Development Fund |
| **ERFS** | Tax and Social Income Survey |
| **ERR** | Expenses-to-revenue ratio |
| **ESF+** | European Social Fund Plus |
| **ESPON** | European Spatial Planning Observation Network |
| **ESSPROS** | European system of integrated social protection statistics |
| **EU** | European Union |
| **FINA** | Financial Agency |
| **GDP** | Gross Domestic Product |
| **GDPR** | General Data Protection Regulation |
| **GIS** | Geographic Information System |
| **HBS** | Household Budget Survey |
| **ID** | Identification |
| **IMD** | Index of Multiple Deprivation |
| **IMSE** | Index of Multiple Social Exclusion |
| **JOPPD** | Izvješće o primicima, porezu na dohodak i prirezu te doprinosima za obvezna osiguranja (receipts, income tax, surtax, and contributions for compulsory insurance form) |
| **LAU** | Local Administrative Units |
| **MJPADT** | Ministry of Justice, Public Administration, and Digital Transformation |
| **MLPSFSP** | Ministry of Labor, Pension System, Family, and Social Policy |
| **MOF** | Ministry of Finance |
| **MOI** | Ministry of Internal Affairs |
| **MPI** | Multidimensional Poverty Index |
| **MRDEUF** | Ministry of Regional Development and EU Funds |
| **NACE** | Nomenclature statistique des activités économiques dans la Communauté européenne (Statistical classification of economic activities in the EU) |

| | |
|---|---|
| **NDS** | National Development Strategy |
| **NRRP** | National Recovery and Resilience Plan |
| **NSI** | National Statistical Institute |
| **NUTS** | Nomenclature of Territorial Units for Statistics |
| **OECD** | Organisation for Economic Co-operation and Development |
| **OIB** | Osobni identifikacijski broj (personal ID number) |
| **PMM** | Predictive mean matching imputation |
| **PPP** | Purchasing Power Parity |
| **RMSE** | Root mean square error |
| **RRF** | Recovery and Resilience Facility |
| **SB-SISBON** | Bank of Slovenia's Client Crediting Rating |
| **SGA** | State Geodetic Administration |
| **SILC** | Statistics on Income and Living Conditions |
| **SMSD** | Severe Material and Social Deprivation |
| **SURS** | Statistical Office of the Republic of Slovenia |
| **TIPSE** | Territorial Dimension of Poverty and Social Exclusion in Europe |
| **TSI** | Technical Support Instrument |
| **UN** | United Nations |
| **UNDP** | United Nations Development Programme |
| **UNECE** | United Nations Economic Commission for Europe |
| **WB** | World Bank |
| **WG** | Working Group |

# 1. INTRODUCTION

## 1.1. BACKGROUND

**The European Union (EU) has consistently acknowledged the critical challenge posed by poverty and social exclusion, formulating comprehensive strategies and dedicating substantial financial resources to address these issues within its member states.** At the Porto Social Summit in 2021, EU leaders, institutions and social partners agreed to put social at the heart of EU policies.[1] They reaffirmed their commitment to the European Pillar of Social Rights (EPSR), which, as part of its Action Plan, aims to bring at least 15 million people out of the risk of poverty and social exclusion by 2030. Achieving this goal requires a holistic approach that addresses people's needs at all stages of life and tackles the root causes of poverty and social exclusion. EU funding is crucial in this strategy, with significant resources made through various channels, including the European Regional Development Fund (ERDF), the European Social Fund Plus (ESF+), the Recovery and Resilience Fund (RRF) and the EU Technical Support Instrument (TSI). The ERDF is dedicated to strengthening economic, social and territorial cohesion. Meanwhile, the ESF+ serves as the EU's primary instrument for promoting employment and social inclusion. The RRF supports Member States to deal with the aftermath of COVID-19 and Russia's invasion of Ukraine. The EU TSI assists Member States in designing and implementing reforms across various policy areas to address the Country Specific Recommendations within the European Semester framework and implement the RRP. Together, these funds form the financial backbone for the EU's effort to create a more inclusive society where all individuals have the opportunity to engage fully in economic, social, and cultural activities.

**While Croatia has made significant progress in reducing poverty over the last decade, pockets of poverty persist.** Between 2011 and 2021, the share of the Croatian population living on less than US$ 6.85 a day at 2017 purchasing power parity prices fell rapidly from 8.1 percent to 1.8 percent bringing it in line with other EU countries at similar income levels (World Bank 2024). Yet, certain demographic groups remain highly vulnerable to poverty and social exclusion. Marginalized communities such as the Roma experience poverty rates that are over fourfold higher than the national average. In 2022, 57.9 percent of Croatian elderly individuals living alone were at risk of poverty and social exclusion, a figure significantly above the EU average of 31.3 percent.[2] Geographically, the capital city of Zagreb and Croatia's northern regions host the majority of the impoverished population (World Bank in 2024). Furthermore, despite the comprehensiveness of the social assistance system, a substantial number of the poor are not adequately covered. A significant portion of the means-tested social assistance is allocated to the Guaranteed Minimum Benefit and child allowances, which collectively extend support to about half of the poorest population. Additionally, the impacts of social transfers (excluding pensions) on poverty reduction in Croatia are not only among the lowest in the EU but are also on the decline.[3, 4]

**Croatia has intensified its commitments to combat poverty and social exclusion, integrating this objective into its strategic planning,** including the National Development Strategy (NDS) until 2030, particularly under the goal of ensuring a healthy, active, and high-quality life. Following the NDS, the Government adopted the National Plan for the Fight against Poverty and Social Exclusion (2021-2027), which sets out priorities and measures for reducing poverty and enhancing the lives of those at risk. A key initiative under this plan is to develop a robust methodology for assessing poverty and social exclusion to inform effective social policies and fulfill national and international reporting obligations. This aligns

1 https://ec.europa.eu/social/main.jsp?langId=en&catId=89&newsId=10004&furtherNews=yes
2 Eurostat 2024 [online code ilc_peps01n]
3 Eurostat 2024 [online code tespm050]
4 European Commission (2023a)

with global commitments like the United Nations (UN) 2030 Agenda for Sustainable Development and the EPSR. Additionally, Croatia's National Plan for Recovery and Resilience[5] (NRRP) as well as the National Reform Programme for 2023[6] address the European Council's recommendations to improve social policies, including the integration of welfare benefits to address country-specific recommendations and enhance poverty reduction, labor market measures, and digital access, with the European Commission monitoring the progress.

**The Ministry of Labor, Pension System, Family, and Social Policy (MLPSFSP) has made significant efforts to monitor poverty and social exclusion at subnational levels, ensuring that EU funds are allocated effectively for related interventions and reforms.** Improved monitoring methods are expected to produce higher quality data on the needs of Croatia's vulnerable groups, informing social policy aimed at reducing poverty and social exclusion. This effort is set to improve the transparency and effectiveness of social policies and is aligned with the NRRP 2021-2026 milestones and targets.[7] Additionally, these efforts are crucial to meeting the enabling conditions (ENC)[8] necessary for the optimal use of EU Funds, including the ESF+, as part of the 2021-2027 cohesion policy. One of the ENCs requires the MLPSFSP to implement an "evidenced based diagnostic of poverty and social exclusion"[9], addressing a wide range of issues from child poverty to access to quality services for vulnerable children, as well as homelessness, spatial segregation, and the needs of vulnerable individuals of all ages.

## 1.2. OBJECTIVES

**This report is the second in a two-part series with an overarching goal of supporting the MLPSFSP to develop a methodology to regularly track poverty and social exclusion at the NUTS 3 or LAU level.** The first report set the groundwork by identifying and assessing the existing administrative data infrastructure relevant to poverty and social exclusion estimates, pinpointing data gaps, and proposing a range of possible approaches to fill these gaps. Based on the agreement with the MLPSFSP and DG REFORM, this second report focuses on leveraging the forthcoming Central Register of the Population—previously known as the Registry of Population, Families, and Households[10]—to develop a register-based methodology for the measurement of poverty and social exclusion at the NUTS 3 or LAU level. The Central Register of the Population (hereafter, "the Population Register") is expected to bridge the data gaps noted in the first report, thereby providing the essential data needed to effectively monitor poverty and social exclusion. As the Population Register is currently under development, this report relies on the latest information obtained from key stakeholders, supplementing any gaps with informed assumptions. Additionally, the report pinpoints challenges related to data and methodology and recommends potential solutions. Specifically, the report encompasses:

---

5 https://commission.europa.eu/business-economy-euro/economic-recovery/recovery-and-resilience-facility/country-pages/croatias-recovery-and-resilience-plan_en

6 Government of the Republic of Croatia (2023).

7 https://planoporavka.gov.hr/dokumenti-113/113

8 Regulation (EU) 2021/1060 of the European Parliament and of the Council of 24 June 2021, Annex IV, Policy Objective 4. "A more social and inclusive Europe implementing the EPSR" https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32021R1060

9 Regulation (EU) 2021/1060 of the European Parliament and of the Council of 24 June 2021, Annex IV, ENC 4.4. "National Strategic policy framework for social inclusion and poverty reduction" https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32021R1060

10 https://vlada.gov.hr/UserDocsImages/Vijesti/2024/Velja%C4%8Da/21%20velja%C4%8Da/Sredisnji_registar_stanovnistva.pdf

- An overview of key concepts of poverty and social exclusion in the EU, and the trend toward tracking poverty and social exclusion using administrative data.
- An examination of the current development of the Population Register, incorporating the latest updates from the Tax Administration of the Ministry of Finance, which is responsible for its development. This covers the timeline for the Population Register's development, databases and variables, and how this Register will fill the data gaps highlighted in the first report of this series.
- A proposed methodology of measuring at-risk-of-poverty indicators based on the Population Register and Tax Income data
- A proposed methodology of measuring social exclusion indicators based on the Population Register and other relevant administrative data sources.
- Options for institutional arrangements that enable data collection, analysis and reporting.
- A proposed monitoring and reporting system for tracking poverty and social exclusion indicators at subnational levels.

## 1.3. METHODOLOGY

**The assessment is conducted through a multifaceted approach that includes desk research, interviews with holders of relevant administrative data, and exchanges with selected National Statistical Institutes (NSIs) across the EU.** The desk research focused on relevant national and EU legislation and regulation, administrative data forms to collect information from individuals and other parties, the websites of institutions responsible for data collection and administration; as well as academic articles and reports utilizing such data. The team held discussions with representatives from the institutions responsible for the collection and management of the administrative data sources, including the Tax Administration of the Ministry of Finance (MOF), the Croatian Bureau of Statistics (CBS), the Ministry of Internal Affairs (MOI), the Ministry of Justice, Public Administration and Digital Transformation (MJPADT), the State Geodetic Administration (SGA), and the Ministry of Regional Development and EU Funds (MRDEUF). In addition, the team exchanged information with selected NSIs to gain insight into their use of administrative income data for poverty measurement; their methodology, if any, to correct for the under-reporting of income; and their use of spatial price differences.

**The report also incorporates insight gained from study visits to Slovenia and Estonia where the Croatian Government Administrative Bodies learned practical experiences from their EU peers.** The two countries were chosen in agreement with the MLPSFSP and DG REFORM for several reasons. Firstly, these two countries have more recent experiences in building a comprehensive administrative data infrastructure compared to the longstanding traditions in the Nordic countries. Additionally, both Slovenia and Estonia are relevant to the country context of Croatia in terms of population size and administrative divisions as a proxy for the scale and complexity of the administrative registries. Finally, Slovenia and Estonia offer a strategic model for Croatia's data development roadmap. Croatia can aim to achieve a system akin to Slovenia's in the medium-term given their shared regulatory and administrative history. In the long-run, Estonia's highly digitalized public sector represents an aspirational goal, reflecting a level of digital maturity that Croatia can strive for in developing and implementing a comprehensive administrative data infrastructure.[11] The study visit to Ljubljana, Slovenia, was organized on January 31st and February 1st, 2024 with Slovenian speakers from the Surveying and Mapping Authority, the Statistical Office, the Ministry of Cohesion and Regional Development, the Ministry of Labour, Family, Social Affairs

11 For reference, in 2022, the EU's Digital Economy and Society Index in Public Services ranks Croatia number 18th out of 27 member states while Estonia and Slovenia were no. 1 and 13, respectively.

and Equal Opportunities, and the Institute of Macroeconomic Analysis and Development. The study visit to Tallinn, Estonia, was organized between June 4th and 6th, 2024 with Estonian speakers from the Statistics Estonia, the Ministry of Interior, the Ministry of Economic Affairs and Communications, the Estonian Tax and Customs Board, and the Ministry of Social Affairs. Annex A1 provides more details of the study tours.

## 1.4. STRUCTURE OF THE REPORT

Following this introduction, Chapter 2 examines the EU's official concepts of poverty and social exclusion, relevant data sources, and the EU's shift from survey-based to administrative-based data. Chapter 3 assesses the development of Croatia's planned new Central Population Register and how this Register will fill the data gaps identified in our previous report. Chapter 4 discusses the challenges associated with measuring poverty using register-based data in Croatia, and proposes potential solutions. Chapter 5 focuses on the measurement of social exclusion. Chapter 6 suggests potential options for setting up institutional arrangements. Chapter 7 recommends a monitoring system to track poverty and social exclusion indicators. Chapter 8 outlines the next steps. The Annexes provide technical details pertaining to various sections of the report.

# 2. OVERVIEW OF KEY CONCEPTS AND DATA SOURCES IN THE EU

## 2.1. THE OFFICIAL CONCEPT OF POVERTY AND SOCIAL EXCLUSION

**The success of this exercise depends, first of all, on a shared understanding of the concepts of poverty and social exclusion, given that these terms may carry different meanings for different individuals.** This chapter explores the official definitions of poverty and social exclusion indicators that Eurostat requires every National Statistical Institute (NSI) of EU member states, including Croatia, to produce and publish. It also discusses the advantages and limitations of such indicators in capturing the spatial distribution of poverty and social exclusion across highly detailed geographical units.

### 2.1.1. At-risk-of poverty (AROP)

**Traditional methods for measuring poverty have commonly focused on income or spending levels.** In many countries, poverty measurements are derived from a minimum threshold of income or spending deemed essential to cover the costs of a basic basket of necessary goods and services. In the EU, the key related indicator is the at-risk-of-poverty (AROP) rate.

**Per Eurostat's definition, the at-risk-of-poverty (AROP) rate is the share of the population whose equivalized disposable household income falls below the AROP threshold.** Such a threshold is set at 60 percent of the national median equivalized disposable income, after social transfers.[12] In turn, Eurostat uses the following definitions for household, household disposable income, and equivalized/adult-equivalent:

- **Household**: A household is a social unit whose members have common arrangements, and "share household expenses or daily needs" in a "shared common residence". A household may consist of either one person living alone, or of a group of people (not necessarily related by kinship) living at the same address with common housekeeping—e.g., sharing at least one meal per day or a living room.
- **Household disposable income:** This is the amount of money available to the household for consumption and saving, after paying taxes on personal income and wealth, social security contributions, and private transfers (i.e., to other households). It is calculated as the sum of wages and salaries, self-employment income, public and private transfers received, and other sources of income (such as capital income) of all members of the household, minus taxes on income and wealth, minus social insurance contributions, minus private transfers that household members pay out (Table 1).
- **Adult equivalent:** To obtain the equivalized disposable income of a household, its disposable income is divided by the number of adult-equivalents using the OECD-modified scale. To calculate the latter, each household member is first classified as either an "adult" (aged 14 or more) or a "child" (aged under 14). Then, the first adult is assigned a weight of 1, every other adult a weight of 0.5, and every child a weight of 0.3. The resulting weighted figures are added up to obtain the number of adult-equivalents in the household.

12 https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Glossary:At-risk-of-poverty_rate

**Table 1.** Income definitions

| Income concepts | Definition |
|---|---|
| **Gross household income** | Total household income before taxes and social security contributions |
| + Wages and salaries | Employee income, including in cash/near-cash and non-cash form (employers' social security contributions are not included) |
| + Self-employment income | Income from self-employment (excluding goods produced for own consumption) |
| + Pensions | Pensions received from an individual private plan |
| + Transfers | Current transfers received (including social benefits and regular inter-household cash transfers received) |
| + Other income sources | E.g., capital income |
| **Tax and private transfers paid** | Tax paid by a household on income, wealth, and social insurance contributions, private transfers sent (employers' social insurance contributions are not included) |
| - Tax on income | |
| - Social security contributions | |
| - Regular taxes on wealth | |
| - Regular inter-household cash transfers paid | |
| **Disposable household income** | Gross household income after subtracting taxes and private transfers paid |

*Source: Atkison et al. (2017).*

## 2.1.2. Social exclusion

**Measures of monetary poverty may not fully capture all dimensions of well-being.** Another approach involves evaluating deprivation across several dimensions, such as living standards, housing, education, and health, among others. Along these lines, the notion of social exclusion encompasses a lack of access to certain key aspects of life, such as the labor market, basic services, and necessary goods.

**Eurostat provides a clear definition of "severe material and social deprivation" (SMSD), to reflect the forced lack of certain items and capabilities that a European needs for a decent life.** In this context, individuals are in a state of deprivation not because they choose to forgo such items/capabilities, but because they cannot afford them, or do not have access to them due to other barriers (e.g., of a social nature). Specifically, people living in SMSD lack at least seven out of 13 so-called deprivation items (Table 2).[13]

13 https://ec.europa.eu/eurostat/statistics-explained/index.php?curid=99141&oldid=534257

**Table 2.** Items/capabilities necessary to avoid severe material and social deprivation, per Eurostat

| Level | Deprivation item |
|---|---|
| **Household** | 1. Capacity to face unexpected expenses<br>2. Capacity to pay for a one-week annual holiday away from home<br>3. Capacity to bear payment arrears (on mortgage or rental payments, utility bills, hire-purchase installments, or other loan payments)<br>4. Capacity to afford a meal with meat, chicken, fish, or vegetarian equivalent every second day<br>5. Ability to keep home adequately warm<br>6. Access to a car/van for personal use<br>7. Capacity to replace worn-out furniture |
| **Individual** | 8. Access to an internet connection<br>9. Capacity to replace worn-out clothes with new ones<br>10. Having two pairs of properly fitting shoes (including a pair of all-weather shoes)<br>11. Spending a small amount of money each week on him/herself<br>12. Having regular leisure activities<br>13. Getting together with friends/family for a drink/meal at least once a month |

*Source: Eurostat 2024.*

### 2.1.3. Very low work intensity

**Recognizing the importance of labor income for an adequate life, Eurostat introduced the concept of "very low work intensity" to capture exclusion from the labor market.** The population with very low work intensity is the share of the population living in households whose working-age members work less than 20 percent of their potential combined working time.[14] In this context, working-age individuals are defined as those aged 18 to 64—excluding students aged 18-24, retired individuals, and individuals over the age of 60 who do not work and whose household's primary income derives from pensions. For households consisting only of persons younger than 18, students younger than 25, or persons aged 65 or older, the level of work intensity is not calculated.[15]

**Working time is calculated on the basis of the total months of employment for working-age members of the household within a year.** For those in part-time employment, working months are adjusted to a full-time equivalent figure—e.g., a person employed on a half-time basis for a full year is counted as having six full-time equivalent months of work. To determine the overall work intensity of a household, the number of months worked by all of its working-age members is added up, and divided by the maximum possible number of months they could have worked in the year considered. A household is deemed to have very low work intensity if the resulting ratio falls below 0.2.

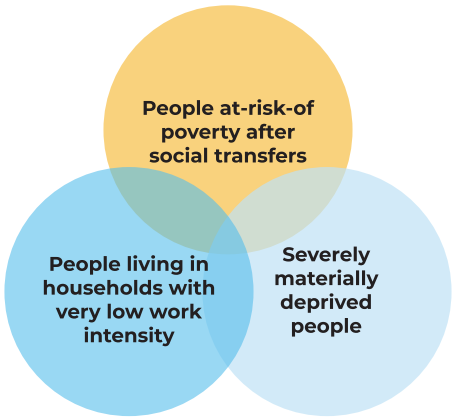
14 https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Glossary:Persons_living_in_households_with_low_work_intensity

15 When determining the AROPE status of a household (see Section 2.1.1), such households are not considered to have very low work intensity.

## 2.1.4. At-risk-of poverty or social exclusion (AROPE)

**The at-risk-of-poverty or social exclusion (AROPE) rate measures the share of the population at risk of poverty, or severely materially deprived, or living in households with very low work intensity** (Figure 1). The notion of AROPE underpins one of the three primary EU targets for 2030, as outlined in the European Pillar of Social Rights. Specifically, the EU aims to reduce the number of people at risk of poverty or social exclusion within the bloc by at least 15 million by 2030.

**Figure 1.** The concept of at risk of poverty or social exclusion, per Eurostat




*Source: World Bank staff elaboration.*

## 2.1.5. Advantages and limitations of the Eurostat framework

**Adhering to the official definitions of poverty and social exclusion allows for methodological rigor and consistency, regular updates to relevant indicators, and cross-sectoral analysis of poverty, social exclusion, and other applicable dimensions.** The NSIs of all EU member states must follow Eurostat's official definitions when producing indicators of poverty and social exclusion—which entails that the methodology for data collection, measurement, and reporting is rigorous and harmonized. The indicators are derived from the annual EU Statistics on Income and Living Conditions (SILC) surveys,[16] and thus they are updated regularly. Moreover, the SILC is a cross-sectional longitudinal survey, addressing a wide range of topics—including household composition, income, employment, education, health, and housing—and tracking changes over time. This comprehensive approach enables a thorough analysis of poverty and social exclusion, which accounts for the intersection of various household characteristics (e.g., gender, educational levels, and health outcomes).

**However, relying on the SILC survey and the AROPE concept limits the geographical granularity and overall scope of poverty measurements.** SILC surveys are generally representative at the NUTS 2 level, and the resulting indicators cannot be disaggregated further. As a result, in 12 out of 27 EU member states,[17] poverty and social exclusion indicators derived from SILC are only available for no more than five regions per country (Table 3)—including in Croatia, where such indicators cover four regions.[18] Such lack of granularity can obscure significant regional and local variations in poverty and social exclusion

16 https://ec.europa.eu/eurostat/web/microdata/european-union-statistics-on-income-and-living-conditions
17 Croatia, Cyprus, Denmark, Estonia, Finland, Ireland, Latvia, Lithuania, Luxembourg, Malta, Slovakia, and Slovenia.
18 Before 2021, Croatia's indicators could only be disaggregated to the level of two regions.

within countries, and can limit the impact of the EU Cohesion Policy. In addition, the concept of AROPE has a narrow scope that mainly focuses on economic, labor, and material deprivation—unlike other frameworks that also consider aspects such as health, education, time, and safety (UNECE, 2022).

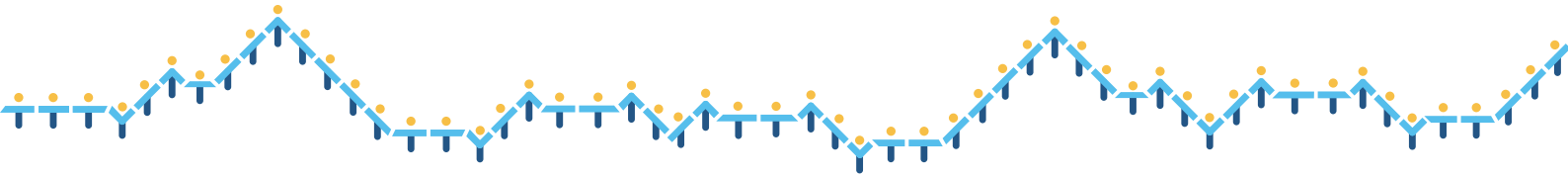
**Table 3.** Geographical disaggregation of EU member states

| | NUTS 1 areas | NUTS 2 areas | NUTS 3 areas | LAU areas |
|---|---|---|---|---|
| **Austria** | 3 | 9 | 35 | 2095 |
| **Belgium** | 3 | 11 | 44 | 581 |
| **Bulgaria** | 2 | 6 | 28 | 265 |
| **Croatia** | 1 | 4 | 21 | 556 |
| **Cyprus** | 1 | 1 | 1 | 615 |
| **Czechia** | 1 | 8 | 14 | 6258 |
| **Denmark** | 1 | 5 | 11 | 99 |
| **Estonia** | 1 | 1 | 5 | 79 |
| **Finland** | 2 | 5 | 19 | 309 |
| **France** | 14 | 27 | 101 | 34956 |
| **Germany** | 16 | 38 | 401 | 10997 |
| **Greece** | 4 | 13 | 52 | 6138 |
| **Hungary** | 3 | 8 | 20 | 3155 |
| **Ireland** | 1 | 3 | 8 | 166 |
| **Italy** | 5 | 21 | 107 | 7904 |
| **Latvia** | 1 | 1 | 6 | 43 |
| **Lithuania** | 1 | 2 | 10 | 60 |
| **Luxembourg** | 1 | 1 | 1 | 102 |
| **Malta** | 1 | 1 | 2 | 68 |
| **Netherlands** | 4 | 12 | 40 | 345 |
| **Poland** | 7 | 17 | 73 | 2477 |
| **Portugal** | 3 | 7 | 25 | 3092 |
| **Romania** | 5 | 8 | 42 | 3181 |
| **Slovakia** | 1 | 4 | 8 | 2927 |
| **Slovenia** | 1 | 2 | 12 | 212 |
| **Spain** | 7 | 19 | 59 | 8131 |
| **Sweden** | 3 | 8 | 21 | 290 |

*Source: World Bank staff compilation.*

## 2.2. MAIN SOURCES OF DATA ON POVERTY AND SOCIAL EXCLUSION

**While the official AROP and AROPE indicators are based on official statistics, EU member states increasingly use complementary administrative data.** This aims to expand the understanding of poverty and social exclusion trends and delve into spatial disparities at finer geographical levels. This section discusses relevant data sources and the ongoing shift from survey-based to administrative-based data in the EU.



### 2.2.1. Official statistics: population census and sample surveys

**The population census is among the most significant activities within the framework of a national statistical system.** As per EU Regulation 763/2008 on population and housing censuses, EU member states must conduct a census at least once every ten years. The latest round of censuses in the EU took place in 2021-2022. The primary purpose of the population census is to gather information about the population and provide a complete and reliable overview of its demographic, social, and economic characteristics, as well as housing conditions. Moreover, the population census serves as the sample frame for topic-specific sample surveys, such as the SILC.

**The SILC is the cornerstone of data collection pertaining to poverty and social exclusion in the EU, with certain limitations.** This annual survey is representative at the NUTS 2 level, providing detailed insight into household income, living conditions, and social participation among the population. It is designed to adhere to internationally recognized standards and methodologies and ensure the comparability of findings across member states. However, the SILC comes with certain notable drawbacks. First, as mentioned previously, its limited geographical disaggregation masks variations at the subregional level. To overcome this, many EU member states have combined the SILC survey with the population census to create high-resolution poverty maps (see section 4.1.1 below), but updating them regularly is a challenge. Second, self-reported data may introduce inaccuracies due to incorrect or partial memories, social desirability bias, and individual interpretations of the questions. Finally, certain population groups may be underrepresented in the SILC; for example, ethnic minorities (such as the Roma) and homeless persons may be difficult to reach or unwilling to participate in the survey, leading to an incomplete picture of poverty and social exclusion.

### 2.2.2. Administrative data and registries

**Certain EU member states have relied on administrative data to obtain timely information on poverty and social exclusion, with a granular level of geographical disaggregation.** As shown in Table 4, a wide range of monetary income and social exclusion indicators are based on data from administrative records—e.g., tax and social security records, which provide insight into earnings, benefits, and social security contributions at the individual level. Administrative data on social transfers—such as social welfare programs, unemployment benefits, and housing assistance—is also key. Employment data, including labor market and payroll records, helps assess work intensity and employment patterns. Furthermore, administrative data enables the monitoring of access to basic services, such as healthcare and education.

**However, to use administrative data effectively, EU member states must establish a robust and well-functioning data ecosystem.** Such an ecosystem encompasses data sources, infrastructure, policies, and stakeholders involved in collecting, processing, and using administrative data. Developing the data ecosystem entails overcoming legal barriers to data access and technical challenges to data integration, ensuring data quality and accuracy, addressing data privacy concerns, fostering a sharing culture among stakeholders, and navigating other legal and regulatory complexities.

**A handful of countries in the EU rely primarily on register-based data for statistical purposes.** Austria, Denmark, Estonia, Finland, Lithuania, Latvia, Malta, the Netherlands, Slovenia, and Sweden have embraced the use of administrative data obtained through ordinary government activities, thereby reducing their dependence on surveys for statistical insight. This model involves a digital data ecosystem built upon existing administrative records and official registers (e.g., the population register, income tax register, building register,

social security register) maintained by government agencies, such as tax authorities and line ministries. Importantly, register-based data offers the potential for cost efficiencies and more comprehensive and accurate information, given its reliance on official records rather than self-reported survey input.

### 2.2.3. The shift from survey-based to register-based data

**The EU statistical regulations have undergone a shift in recent years, with a growing emphasis on using administrative data for official statistics.** This shift is driven by the need for timely and cost-effective statistics that meet the requirements of users and stakeholders. As highlighted in the previous section, a key advantage of leveraging administrative data is the reduction of direct data collection by statistical offices because many relevant data are already collected by various governmental agencies for their internal purposes. In the EU, this shift is evident in multiple statistical regulations where there is a strong encouragement and, in some cases, a mandate for the use of administrative data. For instance, Regulation (EU) 223/2009 on European statistics gives NSIs the right to access administrative data sources for the production of official statistics. In addition, the Integrated European Social Statistics Regulation (EU) 2019/1700 establishes a framework for the integration of various data sources including administrative records. Both regulations aim to reduce the burden on survey respondents while ensuring the delivery of cost-effective, reliable, and high-quality official statistics. Furthermore, the adoption of administrative data opens the possibility of generating statistics for smaller geographical areas, where a survey-based approach would be unfeasible or uneconomical.

**In addition, the official statistics community is moving toward register-based censuses.** The draft EU Regulation on European Statistics on Population and Housing[19] is expected to require member states to submit basic population data to Eurostat on an annual basis, and more detailed information every ten years. The draft regulation will also require member states to establish statistical population registers and use administrative data sources for their maintenance, departing from the traditional field enumeration method. Ten member states—Czechia, Denmark, Estonia, Finland, Lithuania, Latvia, Malta, the Netherlands, Slovenia, and Sweden—implemented population registers based on administrative data that informed their 2021 Population Censuses. Other EU countries have announced plans to follow suit, signaling a broader shift towards register-based censuses for the 2031 round. Although population registers offer crucial information on individuals, many EU administrative systems cannot generate reliable information on households—the reference unit for measuring poverty and social exclusion. In this context, the expected adoption of the draft regulation in 2024 is anticipated to drive significant improvement in administrative systems throughout the EU.

**Moreover, EU member states have made considerable progress in incorporating administrative data in traditional surveys, such as the SILC.** Currently, 20 out of 27 member states pre-fill the SILC's module on income with data from administrative sources. Specifically, nine 'register' countries (Austria, Denmark, Estonia, Finland, Lithuania, Latvia, Netherlands, Slovenia, and Sweden) obtain most income-related data and some demographic information through administrative registers. Eleven countries (Belgium, Bulgaria, Cyprus, France, Hungary, Italy, Ireland, Malta, Spain, and, since 2024, Croatia and Portugal) use a combination of administrative and survey data to construct certain income variables. Only seven member states (Czechia, Germany, Greece, Luxembourg, Poland, Romania, and Slovakia) still rely exclusively on surveys (Table 4). Box 1 summarizes the history of the use of administrative data for the SILC in selected EU countries.

---

19 https://ec.europa.eu/info/law/better-regulation/have-your-say/initiatives/12958-Data-collection-European-statistics-on-population-ESOP-_en

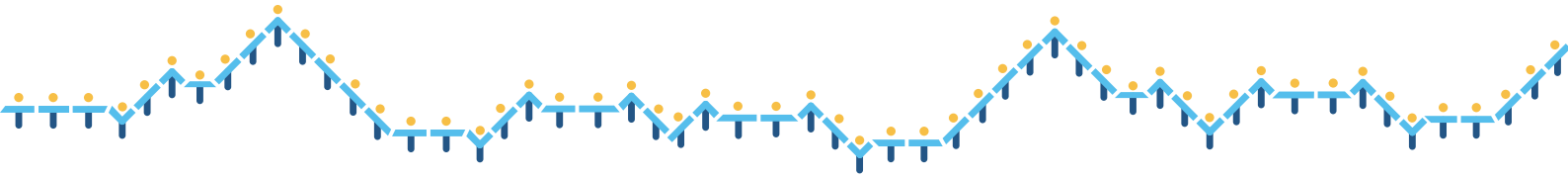


**Table 4.** Use of administrative data and registers in SILC survey, by country

| SILC | EU member states |
|---|---|
| **Sample frame (from population register)** | Austria, Belgium, Denmark, Estonia, Italy, Latvia, Lithuania, Luxembourg, Netherlands, Slovenia, Spain, Sweden |
| **Income** | Austria, Belgium, Bulgaria, Croatia (from 2024), Cyprus (pension and civil servants' income), Denmark, Estonia, Finland, France, Hungary, Ireland, Italy, Latvia, Lithuania, Malta, Netherlands, Portugal (from 2024), Slovenia, Spain, Sweden |
| **Demographic** | Austria, Belgium*, Denmark, Estonia*, Finland, Italy, Lithuania*, Luxembourg*, Netherlands, Slovenia, Spain (partial), Sweden |
| **Education** | Austria, Denmark, Estonia*, Finland, Netherlands, Slovenia, Sweden |
| **Labor** | Austria, Denmark, Finland, Netherlands, Slovenia, Sweden |
| **Housing** | Austria, Denmark, Ireland, Netherlands, Slovenia, Sweden |
| **Countries not using administrative data** | Czechia, Greece, Germany, Luxembourg, Poland, Romania, Slovakia |

*Source: World Bank staff elaboration based on European Commission (2023b).*

** No information in the quality report, but available in the administrative sources.*

# 2. OVERVIEW OF KEY CONCEPTS AND DATA SOURCES IN THE EU

**Box 1.** A brief history of the use of registers and administrative data for the EU-SILC in selected countries

**The ongoing trend in European statistics (including EU-SILC) is to promote the use of administrative data.** At the heart of EU-SILC lies income data, which when gathered through interviews often faces issues like non-responses, leading to the need for imputation to fill in missing observations. Thus, a key advantage of using administrative income data is the signification reduction in such imputations. In addition, income data derived from administrative data are expected to be more accurate and cost-effective to collect, while also substantially reducing the burden on respondents. Nevertheless, the types of errors that affect surveys might also be present in administrative data and registers. Moreover, there can be inconsistencies among different administrative sources, and even among variables within a single source. Table 7 in Section 4.1.2 below will discuss some pros and cons of register-supplemented surveys while this section reviews the experiences of obtaining EU-SILC income data from administrative data in certain EU member states.

**The Nordic countries have used administrative data for statistical purposes since the 1960s.** Their surveys have benefited from a well-established system of basic registers (see Box 2 for more detail). However, the quality of the Population and Dwelling Registers varies across the region (more developed in Denmark, Finland, and Norway, less so in Iceland and Sweden), leading to different approaches to constructing household-level data for the SILC. Sampling modalities also vary, from post-stratified simple random sampling in Iceland to stratified two-phase sampling in Finland, and simple random sampling in Denmark, Norway, and Sweden. The legal basis for access to register data is well established. Technical issues (e.g., record linkage) have mostly been solved.

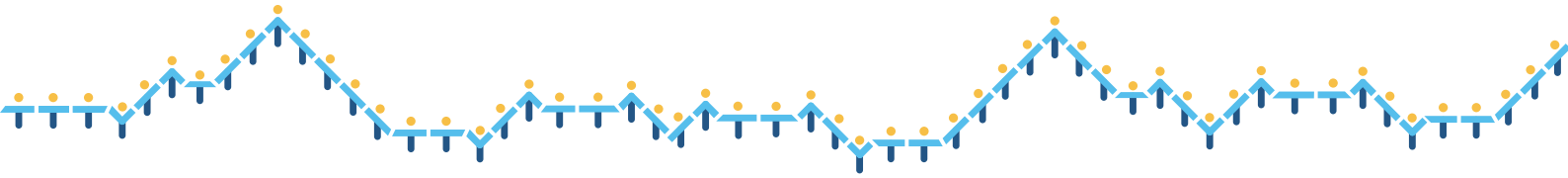
**In France, administrative data has been integrated into survey data since 1956.** Administrative data based income data has been used for the SILC since 2008. A pilot study in 2005 showed that the AROP rates derived from the interview- and administrative data supplemented surveys were similar, at approximately 13 percent each. The shift to administrative based data had a minor impact on the main cross-sectional income poverty indicators, but caused an uptick in the in-come inequality indicator, owing to the broader coverage and greater accuracy of income data.

**In Austria, the switch to administrative data for the SILC was completed in 2012.** The change significantly affected certain income-based indicators—such as income inequality and AROP rates, whose values rose—while the average income figures remained stable.

**In Spain, the income variables of the SILC have been derived from administrative data since 2013.** Spain uses a mixed model, obtaining income data from both registers and questionnaires. Lessons learnt from Spain include the need to first overcome legal barriers. In addition, timeliness is a key constraint, and breaks in time series may arise.

**In Italy, SILC data has been collected from multiple sources since 2004.** The integration of administrative data led to a notable increase in the estimates of average income and the number of self-employed earners, with a smaller rise for employees. At the same time, inequality and poverty rates were revised down.

*Source: Atkinson et al. (2017).*



### 2.2.4. The shift to register-based data in Croatia

**In 2023, Croatia announced its plan to transition to a register-based system, which includes the establishment of a Central Register of Population and conducting a register-based Census in 2031.** The Tax Administration of the MOF made the official announcement of the development of the Central Register of the Population in March 2023. CBS also indicated that the 2021 Population Census would be the final one to primarily rely on field data collection and plan to use the Population Register as the main source for the 2031 Census. As noted in the previous report in this series, Croatia holds extensive administrative data but lacks crucial information to construct household-level data—such as unique identifiers for apartments and information on household members. Chapter 3 details the plan for the Central Register of the Population, which is expected to fill these data gaps and eliminate the need for field data collection in future censuses.

**Croatia recently started using administrative data for certain aspects of the SILC.** Although the CBS had already been tapping administrative sources in various statistical areas, such as wage, employment, and structural business statistics, direct data collection from respondents remained the sole methodology for the EU-SILC until 2023, when individual income indicators were calculated from administrative data for the first time. Specifically, CBS relied on data about income from independent work, pensions (old-age, family, disability, and early retirement), and social benefits collected by the Tax Administration through the JOPPD form. The change of source entailed a break in the time series relative to previous years when data was collected exclusively through surveys.

# 3. CROATIA'S CENTRAL REGISTER OF THE POPULATION: DEVELOPMENT TO DATE

**The recently announced establishment of a comprehensive register of the population and households addresses a major gap in Croatia's register-based framework.** Such a gap was highlighted in the previous report in this series, published in May 2023. In August of the same year, the MOF issued a decision to form and appoint members to a working group (WG) tasked with developing a Population Register.[20] The WG comprises 44 members from 32 government bodies and is led by the MOF and the Tax Administration. After the Tax Administration held bilateral meetings with various government institutions and analyzed their data needs, the MOF presented a draft law on the Population Register in February 2024.[21]

**The previous report in this series noted two key challenges to developing a register-based system in Croatia: the lack of unique identifiers for apartments, and a clear path to determining the composition of households.** Although the exact procedures for addressing both issues are not entirely clear (since the current draft legislation is not accompanied by detailed by-laws), the broad solutions envisaged are presented in Sections 3.4.1 and 3.4.2, while Section 3.7 discusses outstanding challenges.

## 3.1. OBJECTIVES, TIMELINE, AND LEGAL FRAMEWORK

### 3.1.1. Objectives of the Population Register

**The Population Register is envisioned as a centralized data repository that eases the administrative burden in multiple ways.** Specifically, the Population Register aims to:

- Allow for the replacement of the traditional decennial Population Censuses with continuously available data about the population, households, and dwellings.
- Enable a realistic calculation of income per household member to better target social policy and support measures for vulnerable households.
- Make it possible for individuals and households to exercise various rights with no need for requests, supporting documentation, or complex procedures, as the state will have the data to recognize those who have the right to exercise them.
- Enable citizens to view all their personal data in one place.
- Prevent the abuse of social and tax benefits through more up-to-date and interconnected data.

**The MOF has identified 44 specific use cases for the Population Register.** It has especially emphasized the following:

- Maintaining a single electronic work record and a record of inactive persons (to foster the detection of offenders and the suppression of certain forms of undeclared work).
- Speeding up probate procedures, by identifying heirs through data from the Register.
- Analyzing the labor market—i.e., monitoring, assessing, and researching the factors that influence employment trends.
- Calculating the development index and formulating regional development policies.
- Facilitating the enrollment in preschools, schools, student dormitories, and universities (by eliminating the need for applicants to provide supporting documentation).
- Implementing social policy measures (e.g., child benefits, benefits for the elderly, government-funded health insurance, social benefits, and student scholarships).

---

20 https://vlada.gov.hr/UserDocsImages/Vijesti/2024/Velja%C4%8Da/21%20velja%C4%8Da/Sredisnji_registar_stanovnistva.pdf
21 https://esavjetovanja.gov.hr/ECon/MainScreen?entityId=26451

### 3.1.2. Budget and timeline

**The estimated cost of establishing the Population Register and integrating it with other official registers amounts to approximately €12.8 million, allocated from the Government's budget.** In 2023, the development of the Population Register was included in the National Plan for Recovery and Resilience, with a target completion date set for December 2025. The project is expected to encompass four phases:

- **Phase I.** Establishment and implementation of a system for data integration, storage, and analysis (deadline: Q1/2024, value: €2.2 million).
  - Definition of key concepts: population, family, and household.
  - Identification of necessary data and relevant sources.
  - Consolidation of existing data on households in a single repository.
  - Preparation of draft law on the Population Register, with special attention to personal data protection and right to data availability.
- **Phase II.** Establishment of the Register of Population Register Holders (deadline: Q2/2024 – Q3/2025, value: €1.8 million).
- **Phase III.** Establishment of prerequisites for the integration of official records and the SOK Register for 34 state administration bodies (deadline: Q3/2025, value: €8.5 million).
- **Phase IV.** Roll-out of a dedicated portal for citizens within the eTax/mTax system (deadline: Q2/2026, value: €0.2 million)

### 3.1.3. Legal framework

**The draft Law on the Population Register,[22] prepared by the MOF, sets out the legal framework for the establishment and operations of the Population Register.** The draft has passed the public consultation phase and awaits submission to Parliament, with the goal of entering into force starting on January 1st, 2025. However, Articles 29 through 35[23] will be enacted on January 1, 2026, and Articles 13, 24, 25, and 28[24] will come into effect on June 1, 2026. The draft law articulates the Population Register's objectives and scope, provides official definitions for key concepts, and prescribes the key steps for determining the composition of households.

## 3.2. POPULATION COVERAGE, AND KEY DEFINITIONS

### 3.2.1. Population coverage

**The Population Register will cover the entire population of Croatia, encompassing all residents who have an active personal ID (OIB), including those living abroad.** The definition of resident in the draft law is the same as that in the Law on Permanent Residence, whereby a resident is a "Croatian citizen or foreigner with a registered temporary or permanent residence".

---

22 https://esavjetovanja.gov.hr/ECon/MainScreen?entityId=26451

23 Articles 29 to 35 refer to data linking (with other Tax admin records), data usage, publishing of aggregate information, data access, the use of secure data access tools, and the obligation to prescribe the details with a future by-law.

24 Article 13 addresses the voluntary submission of data based on individual declarations. Articles 24 through 25 relate to the declaration process for household members, including those not residing in the same dwelling. Lastly, Article 28 establishes the right to data access for individuals recorded in the Population Register.

### 3.2.2. Key definitions

**The draft Law on the Population Register defines the concepts of resident, kinship, household, and household ID as follows:**

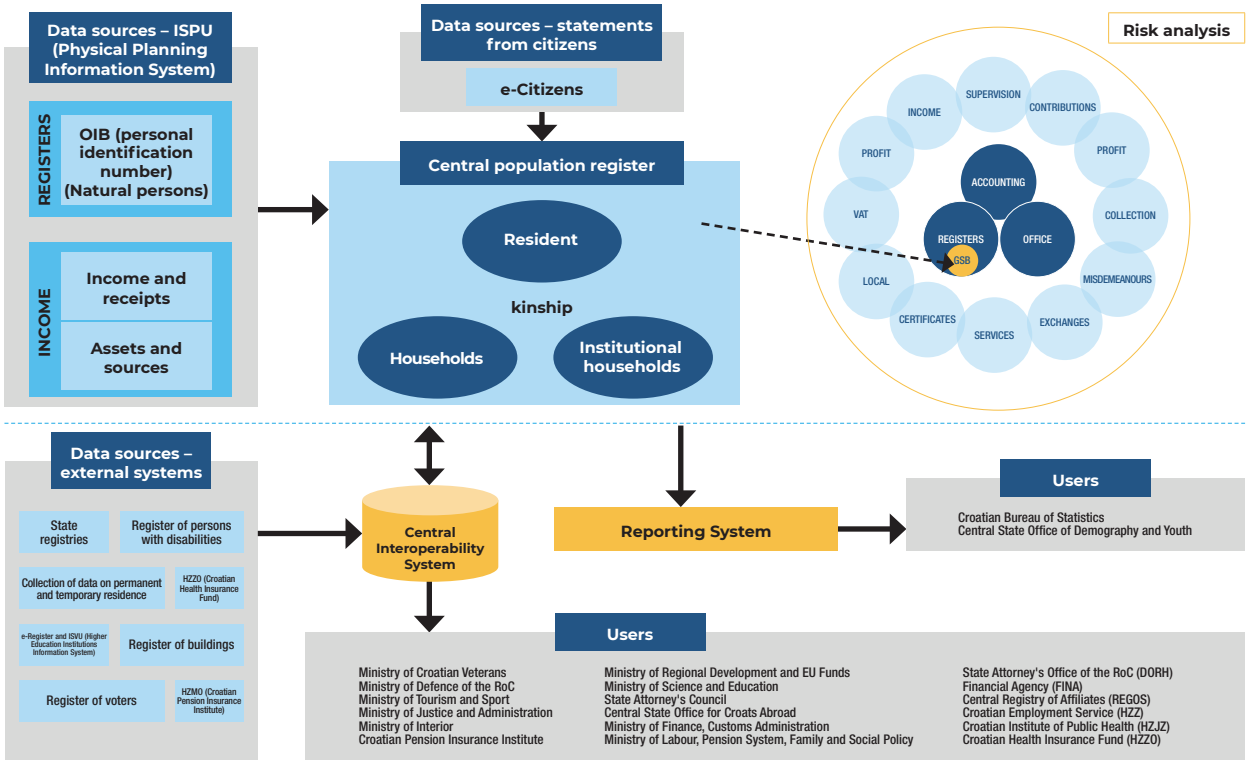
- A **resident** is identified by a unique personal identification number (OIB), the primary piece of data to link personal information across data sources (e.g., information on income, assets, and education). An OIB is issued to all natural and legal persons in Croatia. However, the Register will not solely cover residents, as it will encompass data on:
  - Croatian citizens with permanent and/or temporary residence in the Republic of Croatia, as per residence data registered with the Ministry of Internal Affairs;
  - Croatian citizens with permanent and/or temporary residence outside the Republic of Croatia, as per residence data registered with the Ministry of Internal Affairs;
  - Foreigners holding a long-term, permanent, or temporary residence permit in the Republic of Croatia, as per residence data registered with the Ministry of Internal Affairs; and
  - Persons who are not considered residents according to the provisions of the Law, but are relatives or represent members of the household of registered persons.
- **Kinship** is the connection of a person (uniquely identified through an OIB) with children, a spouse, a common-law partner, a life partner, an informal life partner, parents, grandparents, and siblings, as identified via their OIBs.
- A **household** is a group of persons who live together in a housing unit or part of a housing unit, and who cumulate their earnings and spend jointly. The process for determining the composition of a household is described in Section 3.4.2.
- A **household ID** is the identification number of a residential unit, determined according to a special regulation (see Section 3.4.1). If several households live in the same housing unit, the household ID is the identification mark of the housing unit followed by an ordinal number.

## 3.3. ADMINISTRATIVE REGISTERS USED AS DATA SOURCES

**To ensure comprehensive coverage, the Population Register primarily relies on data from the OIB system.** The OIB is the key personal identifier in all official records in Croatia, and the OIB system already incorporates data on permanent residence for Croatian citizens and temporary residence for foreigners (e.g., registered addresses) from the Ministry of Internal Affairs, as well as regular updates on crucial personal events (e.g., birth, and death) from the MJPADT's Civil Registration System, and other entities.

**Furthermore, the Tax Administration has announced that additional categories of administrative data will feed into the Population Register** (Figure 2). These include data from e-registries from the Ministry of Science and Education; details on the employment status of individuals from the Ministry of Labor, Pension System, Family, and Social Policy; and data on health insurance status from the Ministry of Health. Moreover, data from the State Geodetic Administration's Register of Spatial Units and Register of Buildings will be included in the Population Register. Finally, the Tax Administration will supply income and property data from its Information Systems and connect them with data from the Population Register for analytical and administrative purposes. As the Population Register is still being developed, Box 2 provides an example of the register system in the Nordic countries.

**Figure 2.** Data sources feeding into the forthcoming Population Register




*Source: Tax Administration*

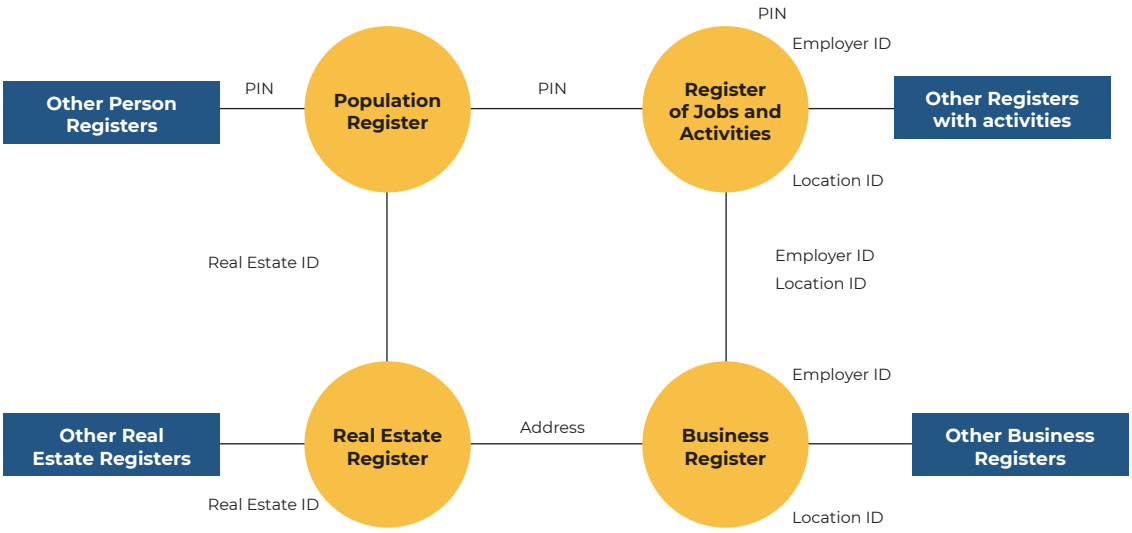
**Box 2.** System of registers in Nordic countries

**The configuration of the register system varies slightly in the Nordic countries.** Information from one or more administrative sources was integrated into the basic registers. Denmark, Finland, Iceland, and Norway have each established three basic registers: a Population Register, a Register of Buildings and Dwellings, and a Business Register. Meanwhile, Sweden has developed a statistical system from four basic registers (Figure 3). This system needs harmonized basic variables, well-defined statistical methodologies, details about mega-data along with guidelines to safeguard privacy (Carlsson and Holmberg 2008). All registers are integrated into the basic registers through the same personal identification number (PIN), business identity number, or in certain cases, real estate ID and location ID (Figure 4).

# 3. CROATIA'S CENTRAL REGISTER OF THE POPULATION: DEVELOPMENT TO DATE

**Figure 3.** Statistics Sweden's system of registers


**Figure 4.** Linkage between Swedish registers




*Source: Jantti et al. (2013).*

## 3.4. FILLING THE DATA GAPS TO IDENTIFY HOUSEHOLDS

**The previous report in this series highlighted that Croatia's administrative data system is currently insufficient for household identification due to the absence of apartment ID.** Eurostat defines a household as one or more individuals living at the same address with common housekeeping.[25] However, Croatia's official address registry lacks apartment numbers for multi-apartment buildings, which is problematic since 91 percent of Croatian residents live in apartments that they own.[26] In addition, the Law on Permanent and Temporary Residence does not include apartment numbers in its definition of "address", and the SGA's Register of Buildings only features apartment IDs for a small number of apartment buildings.

**Other administrative records also fall short of identifying households through kinship.** Kinship-based household identification is also incomplete; the civil registration system, which includes the Birth, Death, Marriage, and Civil Partnership Register, does not cover the entire population. For instance, the Birth Register lacks parental IDs (i.e. OIB or JMBG) for individuals born before 1986, hence excluding 62 percent of the population.[27] Similarly, the Marriage Register does not have spouses' IDs for marriages before 1986, leaving about 17 percent of the population unidentifiable through marital kinship. [28]

### 3.4.1. Identification of apartments

**To address the issue of missing apartment IDs, in 2024, the Ministry of Physical Planning, Construction and State Assets proposed a draft Law on the Management and Maintenance of Buildings.**[29] The draft Law proposes to establish every building community of co-owners as a distinct legal entity, each assigned a unique identification (ID) number by the SGA. This ID would be allocated to every building and its individual units, such as apartments, at the time of the community of co-owners' registration. The draft Law mandates that every building manager submits an officially verified floor plan (or, absent that, a sketch of the floor plan) with data on the usable areas of all specific parts of the building; in this report, we assume that this process will be used to identify apartments. The methodology for assigning unique identifiers to buildings and to each of their separate parts will be prescribed by a future by-law, while the physical markings for every apartment will probably be assigned by building managers. We also assume that the SGA will connect the newly assigned apartment IDs with the apartments recorded in the Land Registry and the Book of Deposited Contracts. These apartment IDs will then become part of the official addresses and be recorded in the SGA's Register of Spatial Units and Register of Buildings. The SGA's methodology for assigning IDs to apartments has not been disclosed; for reference, Box 3 presents examples of approaches taken by the Swedish Tax Agency and the Norwegian Map Authority. Per the current timeline, the Law will enter into force on January 1st, 2025, with an ensuing 12-month period for assigning unique IDs to apartments.

---

25 https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Glossary:Household_-_social_statistics

26 Eurostat (2023). Online code ilc_lvho02

27 World Bank staff's calculation using the 2022 Population census available here https://podaci.dzs.hr/2023/hr/58063

28 This is the upper-bound estimation based on Population Censuses from 1981, 1991, and 2022. The average marriage age in 1981 was 22.3 for women and 25.7 for men. By 1991, these numbers were 23.3 and 27 for women and men, respectively. It is estimated that the average marriage age in 1986 was 24.6 years, calculated as the midpoint between the figures from 1981 and 1991. Thus, it is projected that approximately 17 percent of individuals born before 1961 could have been married in the 2022 census.

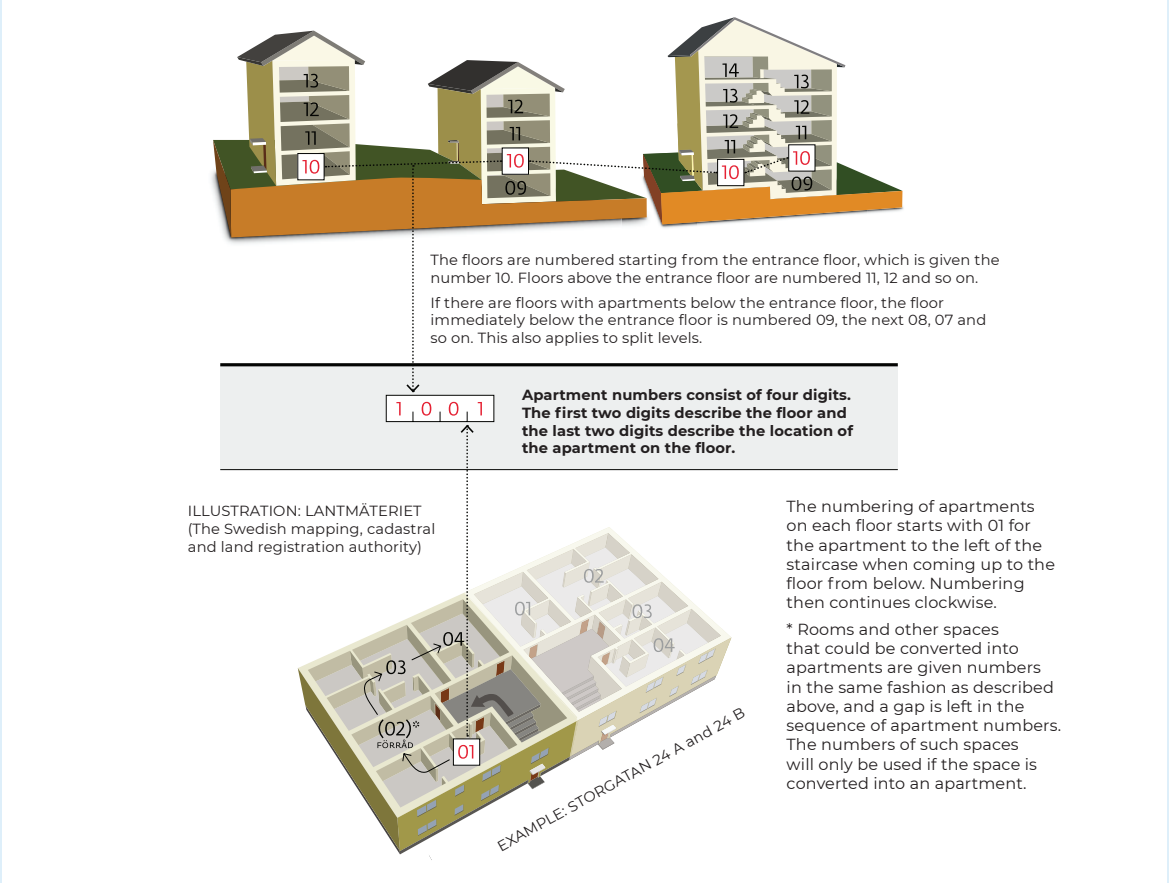
29 https://esavjetovanja.gov.hr/ECon/MainScreen?entityId=26518

# 3. CROATIA'S CENTRAL REGISTER OF THE POPULATION: DEVELOPMENT TO DATE

## Box 3. Identification of apartments: the Swedish and Norwegian approaches

In 2010, the Swedish Tax Agency was tasked with developing and managing a population register at the level of apartments, to enable the collection of statistics on households and apartments. This approach allows cheaper and easier data collection than conducting the traditional population and housing census. The Swedish Tax Agency sent a questionnaire to people registered at the same addresses and people having a family relationship with each other in the population register. The eldest person in each group was asked to fill in the questionnaire, mark their apartment following a harmonized methodology (Figure 5), and submit the questionnaire by mail, online, or by phone.

**Figure 5.** Swedish methodology for apartment numbering




In Norway, the Map Authority requires that all apartments have their address, with a unique apartment number as an identifier. An apartment number consists of a letter and four digits. The letter indicates whether the apartment is on a lower floor (U), basement (K), main floor (H), or top floor (L). The first two numbers indicate the exact floor, and the last two refer to the specific apartment, counted from the left from the perspective of a person who is walking up the stairs. For example, apartment number H0201 indicates the first apartment from the left on the second floor of the Main floor of a building. The Norwegian Tax Administration allows individuals to add or change their apartment number in the National Population Register through an online portal.

*Source: Swedish Tax Agency (2010); The Norwegian Mapping Authority (2023). and The Norwegian Tax Administration (2024)*

### 3.4.2. Identification of households

**Based on the draft Law on the Central Population Register, the identification of households is likely to entail three steps:**[30]

1. The Tax Administration will detect households by identifying kinship among persons who live in the same housing unit (including apartments), relying on existing administrative sources. The search will focus on "immediate family", i.e., couples (by marriage, formal or informal life partnership, or informal union) with or without children (only dependent children will be deemed part of the household, up until their first employment) (Article 22 and 23 of the draft Law).
2. Data that cannot be collected through existing registers will be gathered via individuals' declarations. This will also enable the reporting of larger households, including those comprising people who do not live at the same address, are not related to each other, or may even reside abroad. If a person reports being a member of a particular household, other adult members of that household will be notified and given the possibility to deny the claim within 10 working days. Details of the statement will be outlined in a by-law (Article 24 of the draft Law).
3. Institutional households will be identified by the Tax Administration, through direct contact with relevant institutions (Article 16 of the draft Law).

**At this stage, it is unclear how households can be identified within the timeline of the development of the Population Register.** After apartment IDs are entered into the SGA's Register of Spatial Units and Register of Buildings, the individuals living in each apartment will likely be identified at the time of renewing their personal IDs. The draft Law on the Central Population Register does not require that individuals submit their apartment IDs to the Tax Administration for entry into the Population Register; instead, the draft law appears to envisage gathering such information when individuals renew their ID cards or passports. However, the feasibility of this approach remains unclear at this stage, as it would require a change in the definition of official address to include the apartment number and an amendment to the Law on Residence that defines the addresses down to the house number level. The Population Register and the allocation of unique IDs to apartments are expected to be completed by January 1st, 2026, and an amendment of the Law on Residence is expected in 2024 or in the first half of 2025. Most personal IDs are valid for five years; thus, if linked to their renewal, the identification of apartment residents will be available in 2031. However, a data gap would arise for elderly individuals, as the ID cards issued after August 2nd, 2021, to persons aged 70 and above are valid for 40 years. Therefore, a different mechanism would be necessary to collect data on those persons.

**Other methodologies for connecting individuals to households are possible.** Some of them were discussed in the previous report in this series, including the Slovenian approach based on an administrative census of buildings (see Box 4).

---

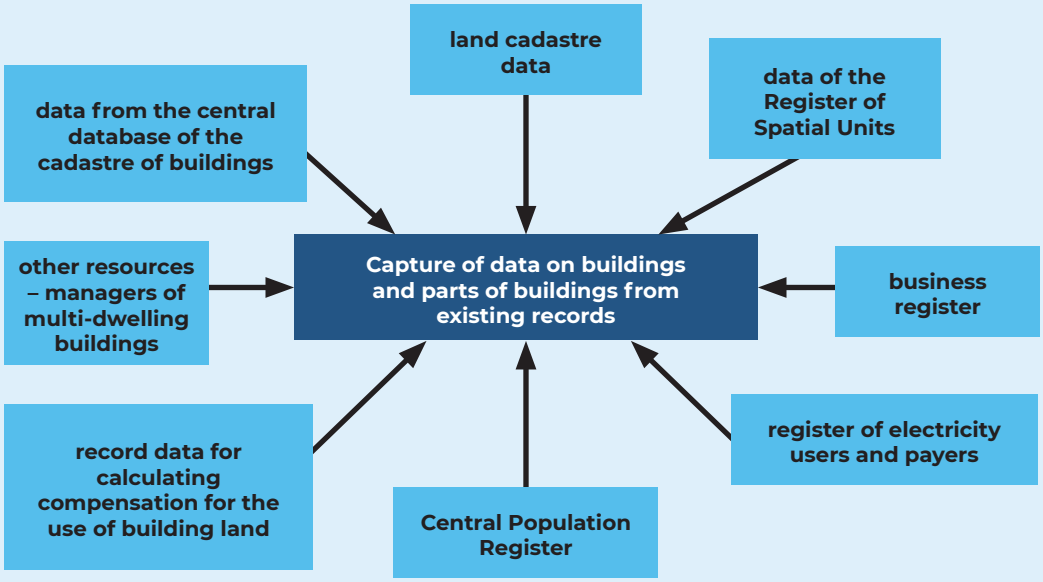
30 The specific rules for generating household information and IDs, and the content of the statement on household members will be determined by the MOF through a by-law.

# 3. CROATIA'S CENTRAL REGISTER OF THE POPULATION: DEVELOPMENT TO DATE

**Box 4.** Slovenia's Building Cadastre and Census of Real Estate

Between 2000 and 2002, the Geodetic Administration of the Republic of Slovenia conducted field identification and photogrammetric capture of 1,401,270 buildings. By 2024, a building cadastre was established by integrating data on buildings and parts of buildings from existing records (Figure 6).

**Figure 6.** Slovenia's Building Cadastre




*Legal basis: Act on the Registration of Real Estate, State Borders and Spatial Units – ZENDMPE (Official Gazette of the Republic of Slovenia, Nos 52/2000, 87/2002 – SPZ and 47/2006 – ZEN*

Between 2006 and 2007, the Geodetic Administration conducted a Census of Real Estate to collect data on housing numbers for all individuals who, at that time, had reported being permanent or temporary residents of Slovenia. The legal basis for this Census is the Real Estate Registration Act, Official Gazette of the Republic of Slovenia, No 47/06, 65/07. Table 5 outlines basic information on the implementation of the Census.

**Table 5.** Slovenia's Census of Real Estate

| Duration of field census | 9 months (1 Dec 2006 – 31 Aug 2007) |
|---|---|
| **Active assistance to owners in record-ing unreported real estate** | Until 21 Dec 2007 |
| **Budget** | €10.9 million |
| **Staff involved** | 1,983<br>• 1,625 inventors<br>• 158 supervisors<br>• 16 State supervisors<br>• 134 assistants |

*Source: Geodetic Administration of the Republic of Slovenia (2024).*

## 3.5. VARIABLES, DATA ENTRY, AND UPDATE

### 3.5.1. Variables and frequency of updates

**Per the draft law, the Population Register will have a general part and a special part, with a list of variables for each part to be set out in a by-law.** The general part will contain the following types of data:

1. Personal Identification Numbers (OIB) and personal names, from the OIB Registry of the Tax Administration;
2. Personal data, from the state register system of the MJPADT;
3. Marriages and life partnerships, from the state register system of the MJPADT;
4. Parents and children, from the state register system of the MJPADT;
5. Legal care and guardianship, from the state register system of the MJPADT;
6. Temporary and permanent residence, from the permanent and temporary residence system of the Ministry of Internal Affairs; and
7. Disability, from the Register of Persons with Disabilities of the Institute of Public Health.

The special part of the Population Register will contain the following types of data:

1. Education, from the Information System of Primary and Secondary Education and the Information System of Records in Higher Education of the Ministry of Science and Education;
2. Data on the basis of health insurance, from the Information System of the Institute for Health Insurance;
3. Data on the basis of pension insurance and employment, and data on pensioners, from the Information System of the Institute for Pension Insurance;
4. Nationality, from the Register of Voters of the Ministry of Administration; and
5. Housing units and housing quality, from the Information System of the State Geodetic Administration.

The special part will also contain the following data, based on individuals' voluntary declaration:

1. Data on extramarital unions or informal life partnerships;
2. Data on religion and mother tongue; and
3. Contact information.

The scope of the data collected will be prescribed in a by-law. In the first meeting of the working group, the Tax Administration provided a tentative list of variables, which will serve as the basis for the by-law. An expanded list of variables, based on information from the draft law, is available in Annex A.2. Notably, the Population Register will include apartment IDs, which at present are not available in administrative registers.

**As the Population Register will integrate multiple administrative data sources, the variables in the general part will be updated as soon as a change occurs in the original databases.** On the other hand, as some variables will only require periodic updates, the data in the special part will be updated once in three months from administrative sources.

### 3.5.2. Data entry, correction, and verification

**Following the integration of existing data sources into the Population Register, the Tax Administration will set up a new online portal.** Such a portal will enable all Croatian citizens who have adequate digital credentials[31] to access their personal data, provide any missing details, and/or correct inaccuracies in the Population Register. The data can be verified when individuals renew their personal documents—such as ID cards and passports—as discussed in Section 3.4.2.

**The new portal is expected to be added to the existing eTax/mTax online platform.** The Tax Administration can consider building this portal upon the success of Croatia's e-citizen system, which has already enabled the collection of reliable household data across a broad segment of the population.[32] Furthermore, it is crucial that the Tax Administration establishes protocols allowing individuals to register their household member(s). The portal is expected to be operational as of June 1st, 2026.

### 3.5.3. Protocols for updating personal information in the Population Register

**According to the draft law, individuals will be able to notify the Tax Administration about any updates or corrections needed in their personal information within the Population Register, however, modifications must be made to the original administrative data sources.** The draft law specifies that the Population Register will entail no changes to the roles and procedures of other registers. Individuals will have access to view their personal information recorded in the Population Register. Should an error in the data be identified by either an individual or the Tax Administration, the Tax Administration will contact the holder of the original administrative source to verify the accuracy of the information. In cases where individuals still have doubts about the data, the Tax Administration will explain which data sources contain the original information and inform the individual of their right to correct any mistake at the source.

## 3.6. POTENTIAL TO MERGE DATA FROM THE POPULATION REGISTER AND OTHER SOURCES

### 3.6.1. Interoperability between the Population Register and other data sources

**For the purpose of measuring poverty and social exclusion, it is important to merge data from the Population Register and other registries, such as the Tax Income database.** The Tax Administration plans to use the pre-existing Electronic Records of Income and Receipts (EDIP) system to verify information on individuals. Potential approaches and institutional arrangements for the calculation of indicators are discussed in Chapter 6. It is important to note that the Population Register will be a basic register, as it will contain at least one authentic piece of information—in this case, information on household members. The data from the Population Register may be shared with other government bodies in compliance with the Law on State Information Structure. This Law, as detailed in the previous report of this series, establishes a framework for the efficient and secure exchange of data. Moreover, the draft Law on the

31 For examples of eligibility, see a list of accepted credential required to access the e-citizen system here: https://gov.hr/hr/lista-prihvacenih-vjerodajnica/1792

32 The e-citizen system has 1,819,779 active users, which comprise 50 percent of the total population and 76.5 percent of the population aged 18-64.

Central Population Register provides a legal foundation for the Tax Administration to merge data from the Register with other data in its Information System for the implementation of its designated tasks. Additionally, the draft Law on the Central Population Register indicates that the provisions of the General Tax Law on tax confidentiality will apply to the Population Register data unless specified otherwise.

### 3.6.2. Legal and technical aspects of data access and sharing

**The draft Law on the Central Population Register defines the general rules and procedures for access to and use of the register's data.** Specifically, the law defines rights of access for registered persons, rights to connect data from the Population Register and from other sources, and rights to use such data for administrative purposes. Rights of access can be broken down into those of the data subject, and those of other parties.

**Per the EU's General Data Protection Regulation (GDPR), a person has the right to obtain information from a processor of Personal Register data about i) their personal data entered in the Population Register, or the absence thereof; ii) the purpose of and legal basis for processing the data; and iii) government bodies and other natural and legal persons who have the right to access data in the Population Register.** Article 28 of the draft Law states that a person recorded in the Population Register has the right to view their registered data, and to request a certificate about it. The GDPR also grants data subjects the right to the erasure of their data; however, Population Register data should fall under one of the exemptions set out in Article 23 of the GDPR, as it is processed to comply with a legal obligation (which should be established by the Law on the Central Population Register). Further restrictions to the rights of data subjects should be considered, in line with the European Data Protection Board's Guidelines on Restrictions under Article 23 of the GDPR.[33] Provisions in the draft law related to data protection will likely be elaborated further after the Croatian Data Protection Authority has commented on it.

**Furthermore, access to the Population Registers and other official records is governed by the GDPR**.[34] More specifically, every governmental entity is required to:

- implement necessary technical, personnel, and organizational measures to protect personal data from accidental loss, destruction, or unauthorized access
- maintain records of individuals who have accessed the requested data
- bear responsibility for any action that violates the relevant legal provisions
- require employees involved in data processing to commit to confidentiality through a signed declaration
- enforce the duty to maintain confidentiality even after the termination of their official duties

**Per the draft Law on the Central Population Register, the Register can be accessed via secure channels, established by the MJPADT.** Registration of data in the Population Register follows the regulations governing State Information Infrastructure, and the register's authentication system must use credentials accepted by the State Information Infrastructure system. The managers of data sources and data users must take measures to ensure secure communication with the Population Register.

---

**33** Guidelines 10/2020 on restrictions under Article 23 GDPR | European Data Protection Board (europa.eu)

**34** Similar provisions for government bodies previously existed in the Law on Personal Data Protection (Official Gazette, No. 103/2003, 118/2006, 41/2008, 130/2011 and 106/2012) which was replaced the GDPR.

## 3.7. OUTSTANDING CHALLENGES

**The process of identifying and assigning apartment IDs appears to overlook smaller structures, potentially impacting older inhabitants.** The draft Law on the Management and Maintenance of Buildings is applicable to structures containing at least four residential units. Nevertheless, the procedure and timeline for assigning apartment IDs to buildings with fewer units remain undefined. Official data on the prevalence of buildings with three or fewer units, as well as the demographics of their occupants, is lacking. However, the MLPSFSP indicates that these smaller buildings are commonly inhabited by the elderly.

**Even for large multi-apartment buildings, the completion timeline can be ambitious.** As discussed in section 3.4.1, the draft Law on the Management and Maintenance of Buildings requires building managers to submit an official floor plan that includes individual apartments. Currently, not every building has an appointed manager. Although there is no official data on the number of buildings without one, as building managers are not currently required to register each building under their supervision, reports indicate that in the city of Osijek alone, there are over 100 buildings without a designated manager.[35] The draft Law also mandates that buildings with four or more apartment units must have a designated building manager, and it outlines a compulsory appointment process by the local authority in cases where co-owners cannot reach a consensus. This process could extend the timeline necessary to complete the assignment of apartment IDs. Although the draft law stipulates penalties for not registering essential details with the SGA, the objective to have every apartment recorded in the building and address registry in all 20 counties and the capital city of Zagreb by January 1st, 2026, appears to be a challenging target.[36]

**The process for mapping apartments to households is not clearly defined, even with the existence of a building registry complete with apartment IDs.** The draft Law on the Central Population Register indicates that family connections from current records will be used initially to establish households, followed by personal declaration when existing data is inadequate. The previous report in this series has highlighted the limitations of using kinship from administrative data sources – in particular the Birth, Death, Marriage, and Civil Partnership Register – to identify households, as these do not cover the entire population. In addition, the draft Law does not specify how the Tax Administration will verify and validate individuals' attestations on their household members. Additionally, it does not outline the process for associating specific apartments with households.

**An alternative approach to identify households is based on place of residence, but the timeline is unlikely to meet the Tax Administration's target.** To identify households through residency, the Law on Residence must be amended to require the inclusion of apartment numbers in official addresses. This amendment is anticipated to occur in the first half of 2025 at the latest. Once the building registry with apartment IDs becomes available on January 1st, 2026, it will be possible to identify the inhabitants of each apartment during their official ID renewals. However, given that Croatian official IDs are typically valid for five years, the entire process could extend until 2031, which overshoots the goal for the Population Register to be operational with apartment IDs and household IDs by the second quarter of 2026. Identifying senior citizens aged 70 and above poses an additional challenge, as their IDs, if issued after August 2nd 2021, have a 40-year validity. Therefore, a separate system will be necessary, for

35 Media reports state that in the city of Osijek alone there were over 100 buildings without an appointed manager.

36 The SGA had a pilot project for data collection for e-building registry in one county. This pilot project took more than 2 years to complete.

example, in 2031, designated staff might need to visit and manually verify the residences of those not yet registered, facilitating their registration on-site.

**The specifics regarding Population Register data access for government entities (apart from the CBS) and researchers for statistical and research purposes are not adequately detailed.** The draft Law on Central Population Register acknowledges the legal rights of certain government bodies to access the Population Register data for their official duties, in line with personal data protection principles and the principles of trust and accuracy, usually on a case-by-case basis (for example, when a Ministry requests personal details of an individual applying for assistance), rather than allowing access to the full database or a random subset. Currently, only the CBS is authorized to use the data for statistical purposes. As the Population Register data is vital for other Ministries, such as the MLPSFSP, for research and operational needs, it is important that the Tax Administration specifies access rules in the Law on Central Population Register, or in a subsequent by-law that will stipulate the conditions, content and the scope of data access for different categories of users. For example, Estonia's Population Register Act[37] grants state and local government agencies the right to access data from the population register for performing public duties. It also allows the release of data for statistical consolidations provided that the data does not contain direct or indirect identifiers. This Act also outlines the procedures for granting access, for instance, through the population register proceedings software, a secure online environment, or an encrypted digital medium.

37 Population Register Act–Riigi Teataja

# 4. REGISTER-BASED POVERTY MEASUREMENTS IN CROATIA

## 4.1. EXISTING AROP MEASURES AND DATA SOURCES IN CROATIA

### 4.1.1. Existing approaches

**The CBS provides the official estimates of the AROP rate in Croatia, based on data from the annual EU-SILC survey.** Such estimates are available at the national (NUTS 1) and regional level (NUTS 2), in line with the survey's representativeness. Since 2021,[38] Croatia has been divided into four NUTS 2 regions: Pannonian Croatia, Adriatic Croatia, City of Zagreb, and North Croatia. The latest AROP rate estimates are based on the EU-SILC 2023[39]—which was the first edition of the survey to use income data partially based on administrative sources (rather than exclusively on self-reporting from respondents), and to adopt the 2021 Population Census as its sampling frame (instead of the 2011 and 2001 Censuses used in earlier years).

**The only available AROP rate estimates below the NUTS 2 level are based on poverty maps produced on a one-off basis by the MRDEUF and the CBS, in collaboration with the WB, in 2016.** Specifically, consumption- and income-based AROP rate estimates were generated using small-area estimation methods that combine survey data (from the 2011 HBS and 2012 EU-SILC) with 2011 Population Census data. The first step in the process entailed estimating the correlation between consumption (or disposable income) and a set of explanatory variables from the surveys. Then, these parameters were used to impute consumption (or disposable income) into Census data to obtain AROP rates at the NUTS 3 (county) and LAU (municipality) levels (MRDEUF 2016). Box 5 highlights other efforts across the EU to produce AROP rates at or below the NUTS 3 level.

**Box 5.** Estimation of AROP rates at or below the NUTS 3 level across the EU

**Many EU countries have collected data to pinpoint pockets of poverty at NUTS 3 level or below, albeit typically on a one-off basis.** Two efforts stand out: the small-area estimation of poverty introduced by the WB; and the maps of the Territorial Dimension of Poverty and Social Exclusion in Europe (TIPSE) project, developed by the Nordic Centre for Spatial Development as part of the EU's ESPON program (European Commission, 2018). The WB methodology was applied to seven EU countries—Estonia, Hungary, Latvia, Poland, Romania, Slovakia, and Slovenia—during the 2014-2020 programming period for the EU's Structural and Investment Funds, in partnership with the EC and the relevant NSIs (Simler, 2016). Subsequently, the initiative was replicated in other EU member states, including Croatia and Bulgaria (Republic of Bulgaria, 2018). In the same period, ESPON's TIPSE project produced maps for all EU member states, plus Switzerland, Norway, Iceland, and Liechtenstein.

38 Before 2021, Croatia only had two NUTS 2 regions: Continental Croatia and Adriatic Croatia.

39 Estimates available at https://podaci.dzs.hr/2024/hr/77038.

### 4.1.2. Relevant administrative data sources

**To measure the AROP rate in Croatia consistently with Eurostat's definition, it is crucial to integrate data from the Population Register and Tax Income.** As outlined in Section 2.1.1., the AROP rate calculation hinges on microdata on household composition, disposable income, and the concept of adult-equivalence. The Population Register is instrumental in identifying households, and establishing adult-equivalence, while providing comprehensive demographic details of Croatian households. It ensures that all segments of the population are represented in the AROP rate estimation, including those who may not be active in the labor market or visible in tax records, such as children. Conversely, Tax Income data in the EDIP system (see Box 6), offers detailed financial information essential for assessing the economic resources available to individuals and households. Merging these two data sources offers a nuanced understanding of socio-economic conditions, helps identify the individuals, households, and groups most vulnerable to poverty, and aids in developing targeted policies to address poverty more effectively. This approach ensures that the measurement of poverty reflects both the population's structure and the distribution of income, in line with Eurostat's poverty measure standards.

**Table 6 lists the data source for each element needed to calculate the AROP rate.** It is also important to note that the Population Register identifies households based on their registered place of residence, which may not align with their actual living arrangements. These disparities can result in skewed household statistics. Box 7 explains how Statistics Estonia addressed a similar challenge.

**Box 6.** The Electronic Records of Income and Receipts (EDIP) data system

**The Tax Administration manages comprehensive income records in the EDIP system, which includes information from:**

- JOPPD forms of receipts, income tax, surtax, and contributions for compulsory insurance;
- Annual income tax returns;
- Decisions on the flat-rate determination for tax on income from self-employment;
- Decisions on renting or leasing movable and immovable property;
- Decisions on the flat-rate determination for tax on income derived from renting apartments, rooms, and beds to travelers and tourists and organizing camps; and
- Acts on asset alienation.

# 4. REGISTER-BASED POVERTY MEASUREMENTS IN CROATIA

**Table 6.** Administrative data sources for poverty measurement in Croatia

| AROP components | Variable | Data source |
|---|---|---|
| **Household** | Household identifier | Population Register |
| **Household disposable income** | Employment income | EDIP |
| | Self-employment income | EDIP |
| | Other income | EDIP |
| | Pensions | EDIP |
| | Transfers received | EDIP |
| | Taxes on income | EDIP |
| | Social security contribution | EDIP |
| | Regular taxes on wealth | EDIP |
| | Regular inter-household cash transfers paid | None |
| **Adult-equivalency** | Household members | Population Register |
| | Age of household members | Population Register |

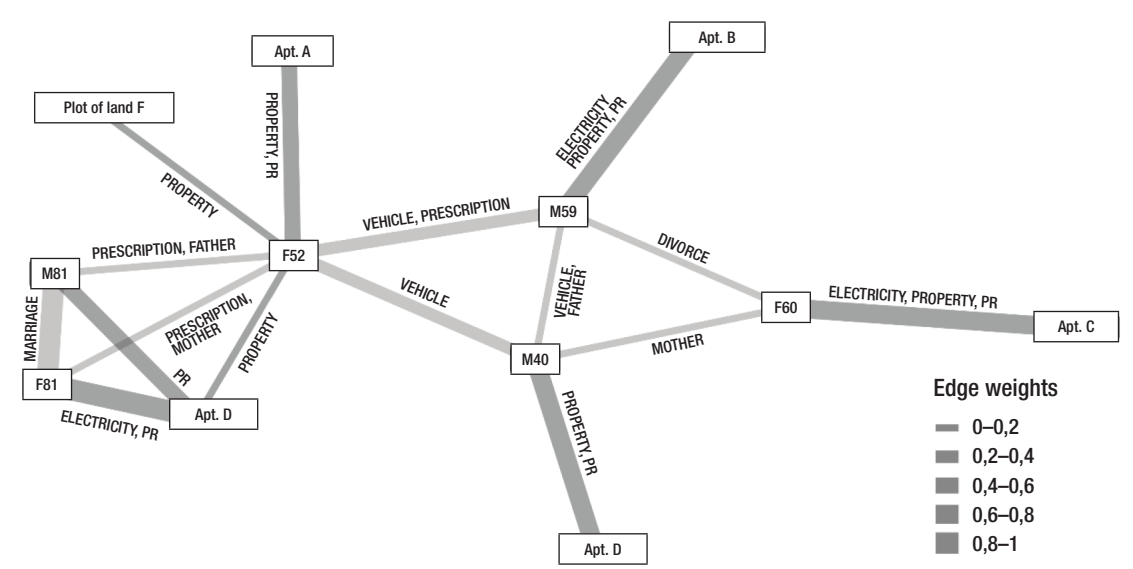
*Source: World Bank staff compilation.*

**Box 7.** Identifying households and dwellings from registers in Estonia

The Estonian Population Register lists an incorrect home address for approximately 20 percent of the population. This is due to a variety of reasons, such as residents seeking to access certain municipal services and benefits, evading tax on rental income, or perceiving registration as unnecessary. Thus, based on data from the Population Register, many households and families appear more fragmented than they are, as their members are listed at separate addresses. In a pilot census conducted by Statistics Estonia in 2016, the composition of households was derived from the Population Register's residence data. The results showed a large deviation from household and family statistics from the 2011 census; for example, the number of single parents was 67 percent higher relative to 2011 figures.

To improve the accuracy of statistics on households and families, Statistics Estonia developed a graph-based methodology that uses data from 17 registers, such as the Population Register, traffic register, and health insurance information system, in addition to data from two surveys—the EU-SILC and the Labor Force Survey. This entails constructing households and dwellings by establishing links between individuals based on relationships such as marriage and parenthood, as well as connections between people and places based on residence, property ownership, and electricity contracts, as shown in Figure 7. Probability models, trained on household data from existing surveys, predict the likelihood of individuals living together or living at a certain address. The resulting household statistics are more consistent with EU-SILC estimates, and mark a considerable improvement over the statistics derived solely from the Population Register.

**Figure 7.** Example of graph of people and places in Estonia




*Note: The nodes are people (with the label showing sex and age) and places (apartments and one plot of land). The edges join people with other people (light grey) or with places (dark grey). Edge weights represent the probability of people living in the same household, or of a person living at a given address. The labels on the edges show the type of relationship.*

*Source: Statistics Estonia (2024).*

**In the medium to long run, the CBS may transition to register-based data collection for the EU-SILC, relying on a range of administrative data sources alongside the Population Register and Tax Income system.** As highlighted in section 2.2.4, Croatia started to incorporate administrative data in the EU-SILC 2023. Using registers offers numerous benefits, such as the elimination of complex and sensitive income-related questions, more accurate data, a lesser burden on respondents, and more efficient and cost-effective surveys (Inglic, 2007). This approach is not without challenges, including the complexities of data integration, the extensive time required for data cleaning, editing, and processing, the issue of unregistered individuals, discrepancies between administrative and statistical definitions, and concerns about data timeliness (Jäntti et al. 2013). Table 7 summarizes the pros and cons of register-based surveys, while Box 8 illustrates the data sources used in Slovenia's EU-SILC.

**Table 7.** Advantages and disadvantages of traditional and register-supplemented surveys

| | Traditional surveys | Register-supplemented surveys |
|---|---|---|
| **Advantages** | • Flexibility of questionnaire.<br>• All data is collected at the same time.<br>• Ability to compare across countries. | • A lesser burden on respondents.<br>• More accurate answers from respondents.<br>• Low cost of data collection.<br>• Almost-full population coverage.<br>• Possibility for reporting on small geographical areas, and regional statistics.<br>• Ability to compare over time (longitude). |
| **Disadvantages** | • Burdensome on respondents.<br>• The quality of answers depends on the respondent's time, willingness to cooperate, memory, and understanding of the question.<br>• Significant cost of data collection.<br>• Low quality of estimates for small geographical areas. | • The inflexibility of information collected (cannot ask or change questions).<br>• Dependent on the administrative data sources' population coverage, objectives, and definitions of variables.<br>• Lack of timeliness, due to mismatch between reporting times for administrative data and for statistics.<br>• Challenges in cross-country comparison, due to differences in administrative data systems.<br>• Low quality of data about variables less important for administrative purposes. |

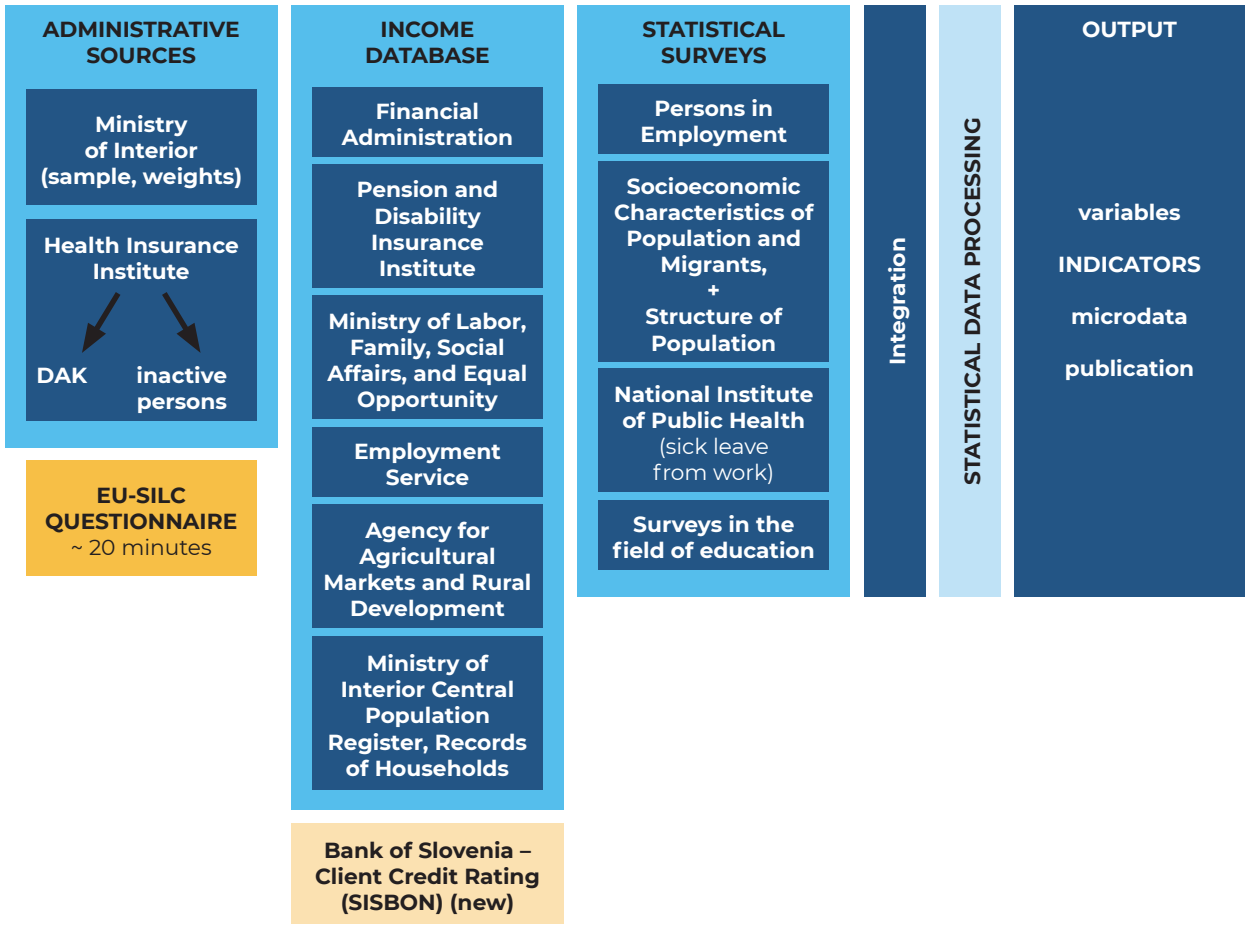
*Source: Wallgren and Wallgren (2014), Jäntti et al. (2013).*

**Box 8.** Data sources for EU-SILC in Slovenia

The Statistical Office of the Republic of Slovenia (SURS) started using registers and administrative data as early as in the 1970s. Today, it is one of nine statistical authorities in the EU[40] that predominantly use registers and administrative sources for the various modules of EU-SILC, including those related to income. Over the years, SURS has progressively integrated additional administrative data sources relevant to EU-SILC, with the Bank of Slovenia's Client Crediting Rating (SB-SISBON) database as the latest addition. As the survey is pre-filled with administrative data, the questionnaire can be kept short, and the interviews last approximately 20 minutes. The main challenge of using administrative sources for the EU-SILC is timeliness, although progress has been made on this front over the years.[41] Figure 8 lists the data sources used for Slovenia's EU-SILC as of 2024.

40 Along with those from Austria, Denmark, Estonia, Finland, Lithuania, Latvia, the Netherlands, and Sweden.
41 Atkinson et al. (2017). Monitoring Social Inclusion in Europe. Eurostat Statistical Book. Luxembourg.

**Figure 8.** Data sources for EU-SILC in Slovenia




*Source: Statistical Office of the Republic of Slovenia (2024).*

# 4.2. CHALLENGES

## 4.2.1. Incomplete or unavailable components in Tax Income data

**This section assesses how the income components of the EU-SILC survey can be calculated from Tax Income data to align as closely as possible to Eurostat's definition.** Table 8 lists the EU-SILC variables used for the construction of household disposable income and evaluates the availability of data in the Tax Income system to replicate these variables. Out of the 20 variables considered in the EU-SILC survey, four are not replicable with Tax Income data, as the relevant information is absent or incomplete.

**Table 8.** Using Tax Income data to construct income components in the EU-SILC survey

| EU-SILC variable | Description | Available in Tax Income Database | Available in other admin. source |
|---|---|---|---|
| **PY010G** | Gross employee cash or near-cash in-come | Yes | - |
| **PY021G** | Company car | Yes | - |
| **PY050G** | Gross cash benefits or losses from self-employment (including royalties) | Partially – for the self-employed paying income tax as a lump sum, only revenues are available; business expenses are not | No |
| **PY080G** | Pensions received from individual private plans (other than those covered by ESSPROS) | Yes | - |
| **PY090G** | Unemployment benefits | Yes | - |
| **PY100G** | Old-age benefits | Yes | - |
| **PY110G** | Survivor benefits | Yes | - |
| **PY120G** | Sickness benefits | Yes | - |
| **PY130G** | Disability benefits | Yes | - |
| **PY140G** | Education-related allowances | Yes | - |
| **HY040G** | Income from the rental of property or land | Partially – for renters of tourist accommodation, only revenues are available; expenses are not | No |
| **HY050G** | Family/children-related allowances | Yes | - |
| **HY060G** | Social allowances not classified else-where | Yes | - |
| **HY070G** | Housing allowances | Yes | - |
| **HY080G** | Regular inter-household cash transfers received | No | No |
| **HY090G** | Interests, dividends, profit from capital investments in unincorporated business | Yes | - |
| **HY110G** | Income received by people under the age of 16 | Yes | - |
| **HY120G** | Regular taxes on wealth | Yes | - |
| **HY130G** | Regular inter-household cash transfers paid | No | No |
| **HY140G** | Tax on income and social insurance con-tributions | Yes | - |

*Source: European Commission (2022) and World Bank elaboration.*

**Tax Income data does not include two EU-SILC income variables: "regular inter-household transfers paid" and "regular inter-household transfers received".** These transfers are often not captured by administrative data sources, as private agreements between individuals remain outside the purview of government agencies. Although certain inter-household payments, such as for child support between divorced parents, could in theory be documented in administrative records (e.g., court orders), obtaining this data can be challenging in practice.

**By excluding inter-household transfers from its calculation, disposable income can be overstated for households that are net payers, and understated for those that are net receivers.** The impact of these inaccuracies on the AROP rate can theoretically be estimated based on the size and distribution of inter-household transfers, using Croatia's EU-SILC survey. However, significant non-response rates in the surveys make it necessary to resort to imputations, potentially compromising the accuracy of such assessments. For example, in the 2021 EU-SILC, the CBS had to impute 22.2 percent of inter-household transfers received and 9.6 percent of transfers paid (Croatian Bureau of Statistics, 2022).

**Tax Income data provides only partial information on two EU-SILC income variables: "gross cash benefits or losses from self-employment (including royalties)" and "income from the rental of property or land".** For the first variable, an accurate calculation requires information about both business revenues and expenses (i.e., the costs incurred to produce goods or provide services). The net result (a profit or a loss) is obtained by subtracting expenses from revenues. However, Tax Income data lacks expense information for self-employed individuals who pay personal income tax as a lump-sum. Under the Law on Personal Income Tax,[42] such taxpayers are only required to report their revenues, which form the basis for their tax liability. On the other hand, for self-employed individuals who pay regular income tax, Tax Income data includes both revenues and expenses. Similarly, "income from rental of property or land" is calculated as the net difference between revenues and expenses, such as property maintenance costs. This category also includes the income of individuals who rent out tourist accommodation, and are subject to a lump-sum personal income tax. As per the Law on Personal Income Tax and the Ordinance on Lump Sum Taxation of the Activities of Renting and Organizing Accommodation in Tourism,[43] their tax liability is calculated based on a fixed rate, determined by the number of beds or camping spots offered and the unit rate established by local authorities. Therefore, these individuals are not required to report their expenses to the Tax Administration. The non-reporting of expenses can cause an overestimation of disposable income for households with members who are self-employed and pay personal income tax as a lump sum, or who rent out tourist accommodation. Consequently, such households can appear less likely to be at risk of poverty than they actually are. The impact of these shortcomings on the AROP rate depends on the size and distribution of unreported expenses.

### 4.2.2. Under-reporting of income in Tax Income data

**Tax Income data is typically sourced from declarations made by the taxpayer or by third parties.** Self-employed individuals are responsible for reporting their own income, while certain parties are mandated to report some financial information pertaining to other parties. Employers, for instance, disclose the wages of their employees, and clients report payments made to contractors under service agreements.

---

**42** Official Gazette No. 115/16, 106/18, 121/19, 32/20, 138/20, 151/22, 114/23.

**43** Official Gazette No. 1/2019.

**Both self-reporting and third-party reporting are susceptible to inaccuracy, including from under-reporting or failure to disclose.** However, under-reporting by a third party might be less likely, as it requires a mutual understanding (whether explicit or implicit) between the taxpayer and the relevant third party (e.g., an employer) (Kleven et al., 2011). Although under-reporting may seem mutually beneficial, the potential gains may not outweigh the risks involved, including the chance of being discovered by tax authorities. Both theory (Kleven et al., 2016) and evidence (Bjorneby et al., 2021) suggest that collusion between employers and employees is more likely in small firms, and unlikely in the public sector (Besim and Jenkins 2005; Paulus 2015).

**Several factors explain the under-reporting of income, including a desire to evade taxes and to remain eligible for certain income-based benefits.** Incentives for tax evasion exist across the income range (Christian 1994; Johns and Slemrod 2010; DeBacker et al. 2020; Auten and Langetieg 2023). Low-income individuals may intentionally report an even lower income to qualify for means-tested social benefits: for example, a person might earn an income in cash in the informal economy, so as to formally remain below the income threshold that qualifies for the Guaranteed Minimum Benefits or for Child Benefits. Establishing the prevalence of this issue in Croatia, however, requires further research.

**Under-reporting extends to earnings from foreign sources.** In 2022, "personal transfers and compensation of employees" or "personal remittances" accounted for 7.6 percent of Croatia's GDP, by far the highest share in the EU.[44] The Law of Personal Income Tax requires Croatian citizens to declare income earned abroad to the Tax Administration, regardless of whether they already paid tax and social insurance contributions in the country of origin. When income is earned in a country that has an agreement with Croatia for the avoidance of double taxation, two approaches to tax liability are possible, depending on the provisions in the agreement. The first approach exempts those who have already paid tax abroad from paying any additional tax in Croatia. The second approach allows for the tax paid in the country where the income was earned to be deducted from the Croatian tax liability. If the amount of foreign tax paid is lower than what is due under Croatian law, the taxpayer must pay the difference in Croatia, which may result in an incentive not to disclose foreign income. The magnitude of such under-reporting is not well-documented. The only available data suggests that from 2004 to 2016, international tax evasion by individuals cost Croatia an average annual loss of tax revenue equal to 0.64 percent of GDP—much higher than the EU average over the same period, equal to 0.46 percent of GDP (Vellutini et al., 2019).

### 4.2.3. Two sets of AROP estimates

**AROP rates derived from registers and administrative data sources are expected to deviate from those calculated from the EU-SILC survey, due to differences in the respective definitions of income, and to behavioral variations observed when responding to surveys versus filing taxes.** EU-SILC surveys employ a broad concept of income, whereas tax data focuses on taxable income. As mentioned in section 4.2.1., Croatia's Tax Income data lacks complete information on four out of the 20 income variables collected in the EU-SILC surveys. Moreover, survey and tax data are influenced differently by factors such as survey response and reporting behavior, tax filing practices, and shifts in the economic, demographic, and legislative context (Bartels and Metzing, 2019). For example, tax filing behavior can be affected by changes in fiscal policies related to income and social benefits—which may introduce bias in the data, particularly around the years of reforms. Moreover, economic growth tends to disproportionately favor

44 The second-highest share was recorded in Latvia, at 3.2 percent of GDP. The EU average stood at 0.8 percent of GDP. https://ec.europa.eu/eurostat/statistics-explained/index.php?title=File:Inflows_of_personal_transfers_and_compensation_of_employees,_2022_(%25_of_GDP).png

high-income earners (Roine et al., 2009), potentially leading to higher values for the inequality indicator when derived from tax data relative to survey data (in which high-income earners are under-represented). Various evaluation studies conducted by NSIs in the EU indicate that the choice between administrative -based and survey-based income data can generate significant differences in the resulting estimates of inequality and poverty (Atkinson et al. 2017, Méndez 2015, Nordberg et al., 2004, Statistics Austria 2014), as discussed in Box 9.

**Box 9.** Differences between household income data from surveys and registers, and implications for poverty estimates

**The differences between household income measurements from surveys and from registers affect the accuracy of poverty estimates, as well as the policy decisions based on them.** Angel et al. (2018) examined the Austrian SILC surveys conducted between 2008 and 2011, assessing both register-based and survey-based income data for the same period and set of households. They found that poverty rates were higher when calculated from register data instead of survey data, mainly due to differences in employment income. Furthermore, both the likelihood and magnitude of under-reporting significantly increase with rising income, in line with findings by Flachaire, Lustig, and Vigorito (2021).

**Differences in poverty estimates between countries, based on whether they rely on administrative or survey data, also arise from the relationship between employment status and risk of poverty.** Lohmann (2011) explores differences in EU-SILC findings between countries based on whether they are calculated from administrative or survey data, which are due to different definitions of ''working.'' The analysis generally found less consistency between information on income from work and on job status in countries that use register-based data. This mismatch can affect the reported poverty rates based on a person's employment status, and subsequently, the perceived difference in poverty risk between those who are employed and those who are not. In a notable subset of countries, conclusions about the impact of employment on reducing the risk of poverty might differ depending on whether the underlying data came from surveys (in which individuals report their own earnings and employment status) or administrative data (where information is collected from official records). This highlights the importance of the data collection approach in accurately assessing the role of employment in mitigating poverty.

*Source: Jäntti et al., 2013.*

**In the short run at least, Croatia will likely have two separate sets of estimates for AROP and inequality at the national (NUTS 1) and regional (NUTS 2) levels, as well as across various demographic groups.** The CBS's annual EU-SILC surveys produce estimates of national and regional AROP rates. Meanwhile, the AROP rates derived from Population Register and Tax Income data can be more granular, covering the national (NUTS 1), regional (NUTS 2), county (NUTS 3), and municipality (LAU) levels. The co-existence of two sets of AROP rates at the national and regional levels may generate confusion among the general public, government entities, and the research community. Moreover, discrepancies may arise between the survey-based and register-based AROP rates attributed to distinct segments of the population, such as children or the elderly, further complicating the understanding and interpretation of poverty rates and trends.

**This issue can be resolved once the CBS adopts the register-based approach to the EU-SILC.** In "register" countries such as Denmark, Finland, Norway, and the Netherlands, many statistics on income as well as indicators of inequality and poverty are derived solely from register-based sources (Jäntti et al. 2013)—which are not constrained by sample size, and can provide detailed longitudinal information and highly disaggregated data for small geographical areas. In Slovenia, AROP rates at the NUTS 3 level can be derived from register-based EU-SILC surveys.[45]

### 4.2.4. Spatial price differences

**Household income data may require adjustments, to reflect cost-of-living variations across regions or areas (e.g., urban versus rural), to ensure that poverty measurements are not biased by spatial price disparities.** Although they may earn a similar income, individuals living in different regions of the country pay different prices for the same goods and services. In developed countries, price disparities for tradable goods tend to be small, because of efficient transportation and integrated distribution networks (Deaton and Zaidi 2002). However, spatial price differences can be large for non-tradable goods and services, such as housing. In addition, regional variations in the non-food price index can be much higher than those in the food-price index (Amendola et al, 2023). Without adjusting for spatial price differences, a national poverty line could over-estimate poverty in areas with lower prices, and vice versa (Ayala et al, 2014; Ferreira et al, 2016).

**However, in practice no universal approach exists to adjust for price differences in poverty estimates.** Numerous studies have explored the use of housing price indices as proxies for local price levels, assuming that prices for other goods do not vary (Early and Olsen 2013, Jolliffe 2006, Moretti 2010, Renwick 2009). Others have attempted to apply regional purchasing power parity (PPP) coefficients, but this approach faces numerous theoretical and empirical challenges, including ensuring that the selected basket of goods is both comparable and representative for each region (Ahmad 2003). Empirical evidence shows that poverty estimates are sensitive to spatial price measurement methods (Ayala et al. 2014, Chen et al. 2020).

**Moreover, in many countries, obtaining reliable and timely sub-national price data is a challenge.** While national Consumer Price Indices (CPIs) are frequently updated and widely recognized, regional CPIs can present significant discrepancies, due to differences in data collection methodologies, the basket of goods and services considered, and the timing of data collection. These inconsistencies can undermine the accuracy of poverty measurements. For example, Deaton and Aten (2014) pointed out that a primary error in the 2005 PPP data affected the cross-region price index, which used a different set of commodities to price goods that were not exclusive to any region. Moreover, the development and implementation of regional CPI measures face numerous logistical and methodological obstacles, such as: the complexity of calculating household spending across regions, to derive accurate price weights; the difficulty in fully capturing all consumer spending categories within the limitations of survey data; and the variation in sample sizes among regions, which can lead to representational issues (Dawber et al., 2019).

45 See, for example, Slovenia's AROP rates in statistical regions: https://www.stat.si/StatWeb/en/News/Index/9624

# 4.3. MEASURING POVERTY USING ADMINISTRATIVE DATA IN EU MEMBER STATES

## 4.3.1. Variation of income components in EU-SILC across countries

**Aligning Tax Income data with the income components of the EU-SILC is a challenge not only for Croatia, but for other EU member states as well.** The methodological complexities of the EU-SILC guidelines often lead to an ambiguous categorization of income sources, making it difficult to compare data across countries (Lynn and Lyberg, 2022). Furthermore, a comparative study of 26 EU-SILC countries[46] by Goedemé and Zardo Trindade in 2020 points out the key issues affecting the comparability of income data—encompassing the differences between administrative records and survey data, different definitions of variables, misclassification and/or absence of certain income components, variations in the level of detail of the data collected, and lack of a uniform approach to collecting data on a net or gross basis. Table 9 details potential comparability challenges within EU-SILC income data from different countries.

**Table 9**. Potential issues for income data comparability in EU-SILC across countries

| | Potential issues for comparability with aggregated income variables | | |
|---|---|---|---|
| **Composite Income variables** | **Deviations from the standard definition (equation to compute the variable)** | **Omission of other income target variables** | **Misallocation or omission of income components** |
| **HY010 (total gross household in-come)** | | Serbia | |
| **HY020 (total disposable household income)** | France and Slovenia | Serbia | |
| **HY022 (total disposable household income before social transfers other than old-age and survivor benefits)** | Spain, France, the Netherlands and Slovenia | Belgium and Serbia | Denmark |
| **HY023 (total disposable household income before social transfers in-cluding old-age and survivor bene-fits)** | Spain, France, the Netherlands and Slovenia | Belgium and Serbia | Denmark |

*Source: Lynn and Lyberg (2022).*

**Deviations from Eurostat's guidelines have been noted in several countries, including France, the Netherlands, Slovenia, and Spain.** France and Slovenia, for instance, calculate income variables using net rather than gross income, although this may not impact comparability if their tax-benefit systems accurately capture net income. Spain, on the other hand, employs net income for certain variables, but has not disclosed any adjustments to improve comparability with other countries. Although the Netherlands has indicated adherence to Eurostat's methods for calculating total disposable household income, it

46 The database is available at: https://timgoedeme.com/tools/metasilc-2015/

diverges from the standard definition of taxes on income and social contributions, and overlooks social transfers in its calculation (Goedemé and Zardo Trindade, 2020).

**In addition, the granularity of income data hinges on the method employed for its collection (i.e., whether it is based on registers or surveys).** Even when adhering to Eurostat guidelines, differences in the level of detail are evident across countries. For example, data on income from interest, dividends, and profits from capital investments in an unincorporated business is denoted as HY090 in the EU-SILC. Austria, Croatia, and Finland each approach the collection of this income component differently. Austria uses a single, broad question in the EU-SILC survey, while Croatia uses two distinct questions, isolating the income derived from interest on savings. Meanwhile, Finland collects this information from registers, enabling the separation of various subcomponents—as shown in Table 10. According to Goedemé and Zardo Trindade (2020), other income variables for which this issue is relevant include housing allowances (HY070), survivor's benefits (PY110), sickness benefits (PY120), disability benefits (PY130), education-related allowances (PY140), income from rental of a property or land (HY040), income received by people aged under 16 (HY110), cash benefits or losses from self-employment (PY050), regular inter-household cash transfer received (HY080), and tax on income and social contributions (HY140).

**Table 10**. Differences between Austria, Croatia, and Finland in the level of detail of data collected to compute income from interest, dividends, and profits from capital investments in an unincorporated business

| Austria (survey) | Croatia (survey) | Finland (Registers) |
|---|---|---|
| • Income from interest, dividends, and earnings from assets | • Income from interest on savings<br>• Investments in securities (i.e., income from dividends, shares in company profits, interest on bonds, and other securities | • Income from dividends as earned income<br>• Undefined capital increase and other capital income from taxation<br>• Income from dividends as capital income<br>• Dividend income from abroad<br>• Interest income from capital of a cooperative<br>• Share of interest in a mutual fund<br>• Interest income taxed at source<br>• Other income from interests<br>• Pensions and other income based on private insurance<br>• Compensation from earned income loss based on private insurance |

*Source: Goedemé and Zardo Trindade (2020).*

**Different countries employ a variety of imputation techniques to calculate income variables.** As detailed by Lynn and Lyberg in 2022, 11 imputation methods are currently in use. The most common are median/mean imputation within specific categories, hot-deck imputation, and regression. It is also common for countries to combine different methods: 17 out of 25 member states[47] apply at least two techniques (Table 11). Box 10 provides an overview of key imputation methods.

47 The total number of member states excludes Denmark, which relies entirely on registers; and Portugal, which uses a gross-to-net micro-simulation model to obtain the net income variable.

**Table 11.** Imputation methods in EU member states

| | Median/Mean imputation within classes | Hot deck imputation | Other | Simple regression imputation | Random regression imputation | Cold deck imputation | Total median/ mean imputation | Micro-simulation model | Multiple imputation | Fractional imputation | Predictive mean matching |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Austria | ● | | | | ● | | ● | | | | |
| Belgium | ● | ● | | ● | | | ● | | | | |
| Bulgaria | ● | | ● | | | | | ● | | | |
| Croatia | ● | ● | | | ● | | | | | ● | ● |
| Cyprus | | | ● | | | | | | | | |
| Czechia | | ● | | | | | | | | | |
| Estonia | ● | ● | ● | ● | | | | | | | |
| Finland | | ● | | | | | | | | | |
| France | | | | ● | ● | | | | | | |
| Germany | | | ● | ● | | | | | | | |
| Greece | | ● | | | | ● | | ● | | | |
| Hungary | | | | ● | | | | | | | |
| Ireland | ● | ● | | | | ● | ● | | | | |
| Italy | | | ● | | | | | ● | ● | | |
| Latvia | | ● | | | | | | | | | |
| Lithuania | ● | | | | | | | | | | |
| Luxembourg | | | | ● | | | | | | | ● |
| Malta | | ● | | ● | | | | | | | |
| Netherlands | ● | | ● | | | | | | | | |
| Poland | | ● | | ● | | | | | | | |
| Romania | | | ● | | | | | | | | |
| Slovenia | ● | ● | | | | ● | | | | | |
| Slovakia | ● | ● | ● | ● | | | | | | | |
| Spain | ● | | | | ● | | | | | | |
| Sweden | | | ● | | | | | | | | |
| **Total number of countries** | **11** | **12** | **9** | **9** | **4** | **3** | **3** | **3** | **1** | **1** | **1** |

*Source: World Bank staff compilation from Lynn and Lyberg (2022) and EU Qualitative Studies.*

**Box 10.** Income imputation techniques

**The simplest imputation techniques include deductive and mean imputation.** Deductive imputation is a method to address missing or inconsistent values in survey data. It involves deducing missing values with near-certainty, often based on the patterns of responses to other questions of the survey.[48] This approach uses logic or an understanding of the relationships between variables and units to fill in missing values. Examples include deriving a value as a function of other values, adopting a value from a related unit, and adopting a value from an earlier time point. Generally, this method is applicable when the true value can be derived with certainty or with a remarkably high probability.

Mean imputation, on the other hand, fills in missing values with the mean of the observed data—but does not maintain the relationships among variables, which are crucial for most research purposes. Mean imputation can cause an underestimation of the actual relationship between variables, and can artificially inflate or deflate correlations, depending on which variable is missing. Additionally, it causes an underestimation of standard errors, because it treats imputed values as real data without accounting for estimation error—potentially leading to Type I errors (Bruch, 2023). Mean imputation within classes takes the standard mean imputation technique a step further. Instead of using the overall average income to impute missing values, it calculates the average income within specific groups (classes) defined by relevant characteristics. As seen in Table 11, this method is the one used most frequently by member states.

**Regression imputation tackles missing income data by leveraging relationships between variables in the survey.** Instead of simply using the average, it builds a statistical model predicting income based on other known characteristics such as age, education, and location. This model then fills in missing income values, potentially leading to more accurate and less biased results.

**In donor imputation, a missing value for a variable is filled with data from a similar respondent, based on shared characteristics.** This approach can extend to filling multiple variables at once, helping maintain their joint distribution intact. Within this methodology, hot-deck imputation uses data from the same dataset for replacement values, unlike cold-deck imputation. The methods vary in how they define and identify similarities between cases. Deterministic hot-deck methods select a single, most-similar donor based on a predefined similarity function. Meanwhile, nearest-neighbor imputation finds a donor by minimizing the distance between the donor and the recipient, using a distance function.

48 UNSD, Handbook on Population and Housing Census Editing Revision 1; Series F No. 82/Rev.1

## 4.3.2. Under-reporting of income

**Adjusting for under-reported income does not appear to be standard practice in EU member states.** As discussed in section 4.2.2., income data from administrative sources may be subject to under-reporting, potentially distorting poverty estimates. However, among the 20 countries that use administrative data for income variables in the EU-SILC, none report correcting for this bias (Table 12).

**Table 12.** Adjusting for under-reported income in administrative sources across the EU

| Country | Income variables in EU-SILC based on administrative sources | Corrections for under-reporting of income to tax authorities |
|---|---|---|
| **Austria** | Yes | No |
| **Belgium** | Yes | No |
| **Bulgaria** | Yes | No |
| **Croatia** | From 2024 | No |
| **Cyprus** | Yes | No |
| **Czech Rep.** | No | |
| **Denmark** | Yes | No |
| **Estonia** | Yes | No |
| **Finland** | Yes | No |
| **France** | Yes | No |
| **Germany** | No | No |
| **Greece** | No | |
| **Hungary** | Yes | No |
| **Ireland** | Yes | No |
| **Italy** | Yes | No |
| **Latvia** | Yes | No |
| **Lithuania** | Yes | No |
| **Luxembourg** | No | |
| **Malta** | Yes | No |
| **Netherlands** | Yes | No |
| **Poland** | No | |
| **Portugal** | From 2024 | No |
| **Romania** | No | |
| **Slovakia** | No | |
| **Slovenia** | Yes | No |
| **Spain** | Yes | No |
| **Sweden** | Yes | No |

*Source: World Bank staff compilation from European Commission (2023b) and bilateral discussion with the respective NSIs.*

**However, since the under-reporting of income also affects survey data, it is unclear whether an adjustment of administrative data is warranted.** Social desirability bias can be a significant source of measurement error for wage and pension income reported in the EU-SILC (Angel et al, 2019). In addition, Bollinger and Taseva (2023) found significant inaccuracy in the reporting of unemployment benefits in the Austrian SILC from 2008 to 2011, which led to underestimating the returns to education, and could have altered the outcomes of job-training program evaluations. Specifically, many respondents reported various combinations of benefits and earnings as "earnings" only. Box 11 summarizes empirical findings about the under-reporting of income in surveys across the EU.

**In addition, surveys generally fail to capture the top of the income distribution, which can affect the AROP thresholds as well as the measurement of inequality indicators.** High-income individuals are less likely to participate in surveys, or may under-report their income, due to concerns about privacy, taxation, or social desirability. In addition, they may be under-represented in survey samples because they are a small proportion of the population. Empirical evidence in the EU and around the world shows that the largest gap between surveys and administrative data concerns the top of the income distribution (Blanchet et al. 2022, Burkhauser et al. 2018, Carranza et al., 2023, Flachaire et al. 2021, Yonzan et al. 2020).

**Box 11.** Under-reporting of income in surveys across the EU

**The prevalence of income under-reporting in surveys, including the EU-SILC, is notable.** Kukk et al. (2020) investigated the scale of income under-reporting by self-employed individuals in 14 EU member states, using the expenditure method and harmonized microdata. Their baseline estimations show that on average, unreported self-employment income ranges from less than 10 percent of actual income in Cyprus and Bulgaria, to more than 30 percent of true income in Greece—suggesting a significant variation in tax compliance across the EU. The study attributes the under-reporting of income by the self-employed to multiple factors, such as the absence of third-party reporting, and the substantial discretion that self-employed individuals have in choosing what to report to tax and statistical authorities.

**Income under-reporting is particularly prevalent among the self-employed and those earning rental income.** Törmälehto (2019) compares income aggregates from the EU-SILC and national accounts, revealing that the largest discrepancies concern property-related and self-employment income. The study points out that such gaps are due to measurement errors and differences in conceptual definitions. However, when the microdata is adjusted to account for these gaps, the impact on the AROP rates is modest. Among countries included in the study, the largest disparities in disposable income as reported in the SILC versus the national accounts were observed in Romania and Greece—where the SILC values were only equal to 33 percent and 50 percent, respectively, of the gross disposable income recorded in the national accounts. As expected, the smallest gaps were found in countries that rely on register data, such as Norway, Denmark, and Sweden, where the SILC's income values were close to 100 percent of the disposable income reported in the national accounts.

In Switzerland, Schmutz (2018) estimates that the share of unreported income among the self-employed stands at around 20 percent of actual income in the baseline scenario. This rate is relatively low compared with other estimates obtained through the same methodology, which range from a minimal 14 percent in Finland to a substantial 62 percent in Estonia. The compiled average of unreported income across all countries and periods covered in this study stands at 27.6 percent of actual income.

In Italy, Albarea et al. (2019) found no statistically significant misreporting of employment income, but did uncover substantial misreporting of income from self-employment and rentals. The study underscores the importance of data quality in tax evasion research, and suggests that the integration of administrative data with survey data could enhance the understanding of tax evasion.

### 4.3.3. Two sets of AROP estimates

**EU member states generally rely on the EU-SILC—whether the income variables derive from administrative data or a survey—for reporting AROP rates at different spatial levels, including NUTS 3.** National and first-tier regional estimates (typically NUTS 1 and NUTS 2, depending on the country's size)[49] are derived from EU-SILC surveys.[50,51] In nine EU member states—Bulgaria, Cyprus,[52] Estonia, Latvia, Luxembourg,[53] Malta, Slovakia, and Slovenia—EU-SILC surveys are also representative at the NUTS 3 level, and can inform the relevant AROP rates at this geographical level. Another five EU member states—Denmark, France, the Netherlands, Spain, and Sweden—use administrative data sources to calculate AROP rates at NUTS 3 level or below (Table 13).

**In Denmark, France, the Netherlands, Spain, and Sweden, which use administrative data to report AROP rates at NUTS 3 level or below, multiple estimates of the AROP rate can arise for the same geographical level.** Each country's situation is explained below:

- **Denmark produces two AROP estimates at the national level:** one is derived from the EU-SILC survey; the other from the Income Statistics Register,[54] which contains administrative data on annual income for both individuals and families.[55] Besides using different data, the two estimates are based on different concepts of income: the SILC does not account for imputed income from housing rentals and for interest on mortgages, whereas the Income Statistics Register does. In addition, SILC uses a definition of household that differs from the definition of family used in the Income Statistics Register.[56]
- **France calculates two AROP rates for most of its NUTS 1 regions, as well as for the broader Metropolitan France.** The national level consists of Metropolitan France (i.e., the area of the country that is geographically in Europe) and five overseas territories;[57] the NUTS 1 level includes 13 regions, plus another region that aggregates the overseas territories. For those 13 NUTS 1 regions, there are two AROP estimates: one derived from the EU-SILC, the other from the FiLoSoFi administrative dataset, which integrates tax and social benefits records.[58,59] Similarly, two estimates are produced for Metropolitan France: one from the FiLoSoFi database, the other from the Tax and Social Income Survey (ERFS),[60,61] with the latter serving as the reference. The ERFS survey, albeit larger and lacking a longitudinal component, uses the same administrative sources for income data as the EU-SILC.[62]

---

49 In Malta, the NUTS 1 and NUTS 2 levels coincide. In Cyprus and Luxembourg, the NUTS 1, NUTS 2, and NUTS 3 levels coincide.
50 In general, the EU-SILC surveys are representative at the NUTS 1 and NUTS 2 levels.
51 These estimates are also published by Eurostat: https://ec.europa.eu/eurostat/databrowser/view/ilc_li41/default/table?lang=en.
52 The whole country is considered a NUTS 3 area. See Table 3.
53 The whole country is considered a NUTS 3 area. See Table 3.
54 See poverty rate calculated from Denmark's Income Statistics Register: https://statbank.dk/ifor12p
55 https://www.dst.dk/en/Statistik/dokumentation/documentationofstatistics/income-statistics
56 https://www.dst.dk/en/Statistik/dokumentation/documentationofstatistics/income-statistics/comparability
57 Guadeloupe, Martinique, French Guiana, La Reunion, and Mayotte.
58 https://www.insee.fr/fr/metadonnees/source/serie/s1172
59 See poverty rates based on FiLoSoFi data: https://www.insee.fr/fr/statistiques/7756729
60 https://www.insee.fr/en/metadonnees/source/serie/s1231
61 See poverty rates based on ERFS data: https://www.insee.fr/fr/statistiques/7710966
62 The ERFS survey predates the EU-SILC; moreover, the latter collects more data on living conditions beyond income poverty, and also produces a longitudinal component (Burricand, 2013).

- **The Netherlands produces two separate poverty measures at the national, NUTS 1, and NUTS 2 levels.** These include the AROP rates obtained from the EU-SILC, and the "low-income" metric[63] from Income Statistics data, which is based on tax records. The methodologies differ in terms of income definitions and equivalence scales. For example, unlike Eurostat, the NSI of the Netherlands excludes child support payments and parental contributions to children not living at home from its concept of income. Moreover, the NSI does not consider households composed of students in its calculation, and uses its own poverty threshold known as the "low-income limit"—which is set slightly above the level at which an individual qualifies for social assistance, and is adjusted for price changes over time.[64]
- **Spain generates two sets of AROP estimates at the NUTS 2 level.** One is derived from administrative tax records,[65] the other from the EU-SILC. Initially, estimates from the administrative dataset were deemed experimental, and did not contribute to official statistics. However, as they gained in reliability, stability, and data quality over time, they were ultimately considered robust enough to be incorporated into official statistics.[66]
- **Sweden maintains two national AROP estimates:** one is based on the EU-SILC, the other is derived from tax records and related administrative sources. Although both estimates utilize the same administrative databases, they differ in the adult-equivalence scale used. The SILC-based AROP estimate follows Eurostat's standard adult-equivalence scale (see Section 2.1.1), while the scale used for the register-based AROP rate accounts for children in shared-residence arrangements when their parents live apart.[67]

---

**63** https://opendata.cbs.nl/statline/#/CBS/nl/dataset/85678NED/table?ts=17146378022 85

**64** https://www.cbs.nl/-/media/_pdf/2023/51/armoede-en-sociale-uitsluiting-2023.pdf

**65** https://www.ine.es/dynt3/inebase/en/index.htm?padre=5650

**66** https://www.ine.es/en/experimental/experimental_en.htm

**67** https://www.statistikdatabasen.scb.se/pxweb/en/ssd/START__HE__HE0110__HE0110F/TabVX1DispInkN/

# 4. REGISTER-BASED POVERTY MEASUREMENTS IN CROATIA

**Table 13.** Data sources for AROP rates at or below NUTS 3 level across the EU

| Country | AROP rate estimates available at NUTS 3 level or below | EU-SILC representative at NUTS 3 level | Dataset for estimates at NUTS 3 level or below |
|---|---|---|---|
| **Austria** | No | No | |
| **Belgium** | No | No | |
| **Bulgaria** | Yes | Yes | EU-SILC |
| **Croatia** | No | No | |
| **Cyprus** | Yes [A] | Yes | EU-SILC |
| **Czech Rep.** | No | No | |
| **Denmark** | Yes [B,C] | No | Administrative |
| **Estonia** | Yes | Yes | EU-SILC |
| **Finland** | No | No | |
| **France** | Yes | No | Administrative |
| **Germany** | No | No | |
| **Greece** | No | No | |
| **Hungary** | No | No | |
| **Ireland** | No | No | |
| **Italy** | No | No | |
| **Latvia** | Yes | Yes | EU-SILC |
| **Lithuania** | No | No | |
| **Luxembourg** | Yes [A] | Yes | EU-SILC |
| **Malta** | Yes | Yes | EU-SILC |
| **Netherlands** | Yes [C] | No | Administrative |
| **Poland** | No | No | |
| **Portugal** | No | No | |
| **Romania** | No | No | |
| **Slovakia** | Yes | Yes | EU-SILC |
| **Slovenia** | Yes | Yes | EU-SILC |
| **Spain** | Yes | No | Administrative |
| **Sweden** | Yes [C] | No | Administrative |

*Source: World Bank staff compilation from bilateral discussions with NSIs and desk research (e.g., Poland: https://bdl.stat.gov.pl/bdl/metadane/cechy/2712; Hungary: https://statinfo.ksh.hu/Statinfo/QueryServlet?ha=OB1119; Portugal: https://www.ine.pt/xportal/xmain?xpid=INE&xpgid=ine_indicadores&indOcorrCod=0009821&contexto=bd&selTab=tab2)*

*Notes:*

*[A] The whole country is considered a NUTS 3 area.*

*[B] Estimates are available at the LAU level.*

*[C] Definition of poverty is not identical to Eurostat's.*

### 4.3.4. Spatial price differences

**Despite the theoretical appeal of the exercise, no EU member state currently accounts for spatial price differences when calculating poverty estimates.** All the NSIs contacted for this review confirmed that their methodologies for calculating AROP rates and other income-based indicators, such as the Gini coefficient, do not account for regional price variations. This can be attributed to two main factors. First, Eurostat's methodological publications do not appear to consider the possibility of adjusting for such variations (see Atkinson, Guio, and Marlier 2017; Guio, Marlier, and Norman 2021). Second, only four out of 27 EU member states—France, Germany, Poland, and Spain—collect CPIs at the regional level. France has CPI data for its metropolitan territory, overseas departments, the Paris region, Corsica, and the metropolitan area excluding Paris and Corsica.[68] Germany reports CPIs for its federal states (NUTS 1 regions).[69] Poland and Spain have regional CPIs for their voivodships[70] and autonomous communities,[71] respectively, which align with NUTS 2 regions. Effectively, four of the EU's five most populous countries have regional CPI data, with Italy being the exception.

## 4.4. RECOMMENDATIONS FOR CROATIA

### 4.4.1. Approach to imputation

**The expenses data that is missing from two income components—"gross cash benefits or losses from self-employment", and "income from rental of property or land"—can be estimated through imputation.** As outlined in Section 4.2.1, Croatia's Tax Income data lacks information on the expenses of self-employed individuals and accommodation renters subject to the lump-sum personal income tax regime. This section introduces a conceptual statistical model that could serve for imputation purposes. The full development of this model would require access to Tax Income microdata from the MOF, and is beyond the scope of this report.

**Out of the common imputation approaches reviewed in Section 4.3.1, regression techniques are the most appropriate in this case.** Regression could be particularly useful for deriving the missing expense data for self-employed individuals and accommodation renters subject to the lump-sum personal income tax regime, starting from the expense data of their peers subject to the regular personal income tax regime. Both groups report their revenue,[72] and additional information such as residence/location and sector of activity (per the NACE classification). The imputation method involves a parametric model that examines the relationship between the expenses-to-revenue ratio (ERR) and predictor variables such as revenues, residence/location, and NACE sector. The parameters estimated based on data from regularly taxed self-employed individuals would then be used to calculate the missing ERRs for the lump-sum-taxed group. Expenses are then computed by multiplying the imputed ERRs by the corresponding revenues. It is important to consider the random error in this model, and the confidence intervals of

---

**68** Available here: https://www.insee.fr/en/metadonnees/source/operation/s2124/publications.

**69** Available here: https://www-genesis.destatis.de/genesis/online?operation=table&code=61111-0010&bypass=true&levelindex=0&levelid=1714654225189#abreadcrumb.

**70** Available here: https://stat.gov.pl/en/topics/prices-trade/price-indices/consumer-price-indices-by-voivodships-in-the-fourth-quarter-of-2021,12,17.html.

**71** Available here: https://www.ine.es/jaxiT3/Tabla.htm?t=50913.

**72** Self-employed individuals subject to the lump-sum tax regime report their revenues via the PO-SD forms (https://www.porezna-uprava.hr/HR_obrasci/Documents/POREZ%20NA%20DOHODAK/PO-SD.pdf) and PO-SD-Z forms (https://www.porezna-uprava.hr/HR_obrasci/Documents/POREZ%20NA%20DOHODAK/PO-SD-Z.pdf). Accommodation renters subject to the lump-sum tax regime report their revenues via the EP forms (https://www.porezna-uprava.hr/HR_obrasci/Documents/POREZ%20NA%20DOHODAK/EP.pdf) and TZ-2 forms (https://www.porezna-uprava.hr/HR_obrasci/Documents/OSTALO/TZ%202.pdf)

the estimated parameters. The missing expenses can be imputed multiple times, resulting in a set of *M* imputed values for each individual whose expenses are missing. These various sets of imputed expenses are then used to construct *M* amounts of household disposable income. The AROP rate is estimated *M* times, with the average used as the final estimate. The technical details of the imputation approach and the estimation of the AROP rate from imputed data are outlined in Annexes A3 and A4.

**The two missing income components—"inter-household transfers paid" and "inter-household transfers received"—can be excluded from the estimations of the AROP rate and other income-based measures, provided that this exclusion is clearly documented.** As discussed in section 4.2.1, these items are typically unavailable in or difficult to extract from Croatia's administrative records. To fully align with Eurostat's definition of income, a survey should be conducted for collecting data on inter-household transfers. For example, in Slovenia, the Statistical Office retrieves most EU-SILC income components from administrative databases—with the exception of inter-household transfers, for which data is obtained via a short annual EU-SILC survey.[73] In the long run, when CBS moves toward the register-based EU-SILC approach adopted by many EU member states, it might consider a light EU-SILC survey to complement administrative data. In the interim, for the purpose of computing AROP rates and other income-based indicators used by the MLPSFSP, inter-household transfers can be disregarded. However, it is crucial for the MLPSFSP to document that income measurements will differ from Eurostat's due to this exclusion. Deviations from Eurostat's income concepts already occur in other EU Member States, such as Denmark, the Netherlands, and Sweden, as shown in Table 13.

### 4.4.2. Approach to under-reporting of income

**In line with common practice among EU Member States, it is recommended not to adjust Tax Income data for potential under-reporting.** As discussed in Section 4.3.2, income data derived from both surveys and administrative records is subject to various degrees of bias. The necessity of applying corrections specifically to tax income data is not warranted. Furthermore, the practice of not adjusting for under-reporting bias in tax income data is common across the EU (see Table 12).

**Nevertheless, it is useful to estimate the degree of under-reporting in Tax Income data to understand its potential impacts on estimates of the AROP rate and other income-based indicators.** Although the "true" income distribution is unobservable, the unreported income can be approximated, using survey data to construct a proxy for the actual income distribution from survey data (Pissarides and Weber, 1989). One approach entails classifying households in survey data as either employed or self-employed, with the assumption that the former fully report their income while the latter do not. The extent of under-reporting can be estimated by comparing income and expenses—assuming that, if self-employed households have similar expenses to employed households but lower reported income, the gap indicates under-reporting. This model accounts for additional factors, such as wealth and household composition (see Annex A5 for technical details). The only available estimate of under-reporting in Croatia, based on HBS 2010 data, suggests that self-employed households under-report their income by an average of 16.8 percent; however, this estimate is not statistically significant (Kukk et al., 2020). This analysis could be refined using more recent data, particularly from the HBS 2019 and HBS 2022.

---

**73** The Statistical Office of the Republic of Slovenia (2024). Materials prepared for the Study Visit to Slovenia – Using Data for Poverty and Social Exclusion Measurement and Related Policies, January 31 – Feb 1, 2024, Ljubljana, Slovenia.

### 4.4.3. Two sets of AROP estimates

**The SILC-based AROP estimates at the national and regional levels (NUTS 1 and NUTS 2) should be recognized as the exclusive official figures.** The CBS—which complies with Eurostat's protocols for data gathering, processing, and dissemination—is responsible for these measures,[74] which are available to the public on the databases of both the CBS and Eurostat.[75]

**The register-based AROP estimates at all geographical levels (NUTS 1, NUTS 2, NUTS 3, and LAU) could be labeled as experimental.** It is important to meticulously document and publicly disclose the methodological differences between the AROP indicators derived from Population Register and Tax Income data, and those derived from the EU-SILC, in line with current practice in other EU member states such as Denmark, France, the Netherlands, and Sweden. The potential for such experimental statistics to be recognized as official depends on the institutional frameworks for their production and validation, as outlined in Chapter 6. If the CBS were to produce and validate these statistics, they could achieve official status, as occurred in Spain (see section 4.3.3). If the methodological differences between them are clearly documented, disclosed, and communicated, the co-existence of two sets of AROP estimates for the same geographical area is acceptable. As noted in section 4.3.3, dual sets of estimates are produced in five EU member states that use administrative data for generating AROP rates at NUTS 3 level or below (i.e., Denmark, France, the Netherlands, Spain, and Sweden).

### 4.4.4. Spatial price differences

**To align with the approach of other EU member states and reflect the lack of regional CPIs, it is recommended that Croatia does not account for spatial price differences in the context of poverty measurements.** As discussed in Section 4.3.4, all NSIs across the EU confirmed that they do not account for such price variations. Moreover, in Croatia, the CBS only produces a national CPI. The generation of CPIs at any sub-national level, including for the four NUTS 2 regions, is not feasible at this stage for two main reasons. First, the HBS—which is instrumental to the construction of the reference consumption basket—is only representative at the national level. Second, the collection of price data is currently limited to nine cities (Zagreb, Slavonski Brod, Osijek, Sisak, Rijeka, Pula, Split, Dubrovnik, and Varaždin), which precludes the calculation of regional CPIs.

---

**74** See Croatia quality report for EU-SILC: https://circabc.europa.eu/ui/group/853b48e6-a00f-4d22-87db-c40bafd0161d/library/011ca788-7678-42fe-bc74-f68b087257bc

**75** https://ec.europa.eu/eurostat/databrowser/view/ilc_li02/default/table?lang=en

# 5. REGISTER-BASED MEASUREMENTS OF SOCIAL EXCLUSION IN CROATIA

## 5.1. CONCEPTS OF SOCIAL EXCLUSION

**The notion of "social exclusion" is multifaceted, and does not have a single, universally accepted definition.** The concept emerged in the late 1970s, originally to acknowledge and address the issue of marginalization in French society. However, the lack of a precise definition of social exclusion makes it challenging to measure its extent, or to quantify the number of individuals it affects. Unlike the notion of poverty, which focuses on an outcome, social exclusion emphasizes both outcomes and the process that leads to them—i.e., to the exclusion of individuals and groups from society (UNECE, 2022).

**Measuring social exclusion offers a deeper understanding of disadvantaged groups, beyond a focus on monetary poverty.** Empirical evidence shows that social exclusion can reduce pro-social behavior, highlighting the significance of fostering inclusion to promote community support and cooperation (Twenge et al., 2007). Furthermore, social exclusion can have detrimental effects on mental health and the capacity of individuals to thrive in society (North and Fiske, 2013). Therefore, measuring social exclusion is crucial for policymakers to identify individuals and communities that are systematically marginalized, and to understand the mechanisms of their exclusion so as to design effective interventions. Nowadays, social exclusion indicators are incorporated into national and international policy efforts. For example, Australia[76] and Canada[77] have adopted policy measures to assess and improve social inclusion. Moreover, social inclusion is a key element in the Sustainable Development Goals for 2030 and the Europe 2030 strategy.

**Despite the widespread acknowledgment of its importance, there is no harmonized approach to measuring social exclusion.** The assessment of social exclusion or inclusion generally focuses on specific domains or aspects of life that are linked to social exclusion theories, policy objectives, or both. Various indicators capture different facets of each domain. The domains include, but are not limited to, income-based poverty, exclusion from the labor market, education and skills, health and disabilities, and access to public services. Box 12 summarizes social exclusion frameworks used in high-income countries across the world. Table 14 presents key domains, and the frameworks in which they are applied.

**Box 12.** Frameworks of social exclusion, and examples from EU member states and high-income countries

**EU AROPE indicator:** The at-risk-of-poverty or social exclusion (AROPE) indicator is recommended by Eurostat to measure the poverty and social exclusion targets associated with the Europe 2030 Strategy. The strategy's objective is to deliver inclusive growth and lift at least 15 million people across the EU, including at least 5 million children, out of poverty and social exclusion between 2019 and 2030.[78] The AROPE concept is explained in more detail in section 2.1.4. Albania and Romania also use it as principal measure of social exclusion.

**Multi-dimensional poverty indices:** A Multidimensional Poverty Index (MPI) is used to observe the proportion of the population that could be poor with reference to various aspects of life (Alkire and Foster, 2011). The indicators chosen can be adapted to meet the needs and policy focuses of

76 Australian Social Inclusion Board (2012). Social Inclusion in Australia: How Australia is faring. Department of the Prime Minister and Cabinet.

77 https://www.canada.ca/en/employment-social-development/programs/poverty-reduction/reports/strategy.html

78 https://ec.europa.eu/commission/presscorner/detail/en/ip_22_3782

each country. Numerous countries are establishing National MPIs as part of official statistics on poverty. Certain high-income countries, such as Chile and the US, already use them, and Germany is contemplating adopting one. In Chile, for example, the the MPI encompasses education, health, employment and social security, and basic standards of living.

**Well-being framework and indices:** The OECD Well-being Framework (The Better Life Initiative) encompasses 11 dimensions that capture crucial elements of present well-being, and four dimensions relevant to future well-being. Indicators of current well-being comprise income and wealth, job quality and employment, housing conditions, health status, knowledge and skills, environmental quality, personal well-being, safety, work-life balance, social connections, and civic engagement. The dimensions of future well-being include natural capital, human capital, social capital, and economic capital. For example, Italy employs a multidimensional method to measure "equitable and sustainable well-being", capturing indicators of well-being, inequality, and sustainability alongside traditional indicators of production and economic activity. A total of 130 indicators are utilized across 12 domains of well-being: economic well-being; education and training; environment; health; innovation, research, and creativity; landscape and cultural heritage; politics and institutions; quality of services; safety; social relationships; subjective well-being; and work-life balance.

**Social capital and social cohesion framework:** The notion of social capital captures the value of social networks for economic prosperity and well-being, illustrating how behaviors, attitudes, and interpersonal relationships enhance various aspects of an individual's life. It encompasses principles such as trust, safety, and a sense of belonging. The benefits of social capital can be personal (e.g., family support) or community-wide (e.g., volunteer work), and the level of social capital is linked to economic growth, sustainability, and overall well-being. For example, the United Kingdom focuses on four domains of social capital: personal relationships, social network supports, civic engagement and trust, and cooperative norms. Such domains are covered by 25 indicators, with most of the relevant data supplied by a range of existing surveys. The Netherlands measures both social capital and social cohesion. The former comprises two dimensions: participation and trust (Van Beuningen and Schmeets 2013). In turn, each dimension is broken down into three sub-dimensions: social, organizational, and political. Social participation encompasses meaningful social contact; organizational participation includes membership in organizations, attendance of events, and participation in the labor force or education; political participation covers voting, party membership, and involvement in political action. Finally, social trust refers to forming positive ties with others; organizational trust encompasses trust in institutions, of various types; and political trust concerns trust in political institutions specifically.

**Frameworks applicable to groups at particular risk of disadvantage:** Measurement efforts focusing on groups at high risk of disadvantage are crucial for addressing issues of social exclusion and inclusion, as well as broader well-being. Although such groups are often covered by broader measurements, a comprehensive understanding of their difficulties and the identification of effective solutions may necessitate tailored approaches. Notable examples include the Child and Youth Well-being Strategy in New Zealand, and the multi-dimensional index for measuring poverty and deprivation among Roma people in the Western Balkans.

**For instance, in 2019, New Zealand introduced the Child and Youth Well-being Strategy, a cohesive framework to assess the well-being of children and define what constitutes a good life for them.** It is updated every three years, to remain relevant and responsive to societal shifts. The framework encompasses six interconnected domains, to assess whether: children are loved,

# 5. REGISTER-BASED MEASUREMENTS OF SOCIAL EXCLUSION IN CROATIA

safe, and nurtured; have what they need; are happy and healthy; are learning and developing; are accepted, respected, and connected; and are involved and empowered.

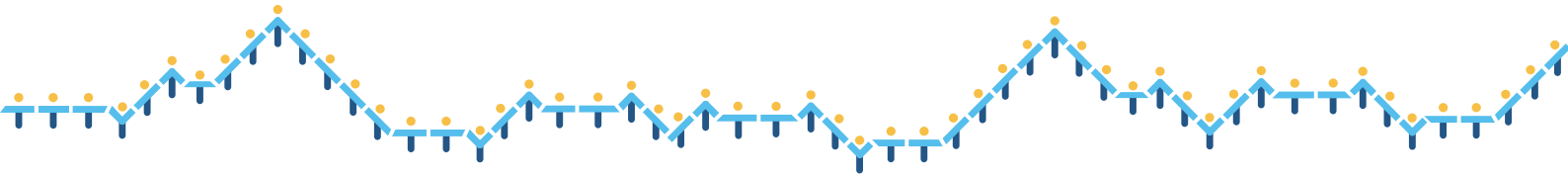
**Another example is the UNDP multi-dimensional index developed for measuring poverty and deprivation among the Roma people in the Western Balkans, and facilitating more targeted policy interventions.** Data for the index came from the Regional Roma Surveys conducted in 2011 and 2017, which offered insights into the experiences of Roma communities in the region. A customized survey allowed for an in-depth exploration of issues relevant to the Roma, such as restricted opportunities for political participation, limited access to social and medical services, and challenges in obtaining personal identification documents. The index incorporates 12 equally weighted indicators across six essential dimensions: basic rights, health, education, housing, standards of living, and employment.

*Source: UNECE (2022).*

**Table 14.** Domains and frameworks of social exclusion

| Domain | Frameworks to which domain is relevant |
|---|---|
| **Income poverty** | EU AROPE indicator; EU Laeken; German well-being indicators; UNDP social exclusion framework (2012); Switzerland's social exclusion framework; Canada's Community Well-being Index; Multidimensional Inequality Framework; Individual Deprivation Measure (International Women's Development Agency); Mexico's National Survey on the Dynamics of Household Relationships (ENDIREH) |
| **Labor market (exclusion or inclusion with regard to participation in the labor market)** | Social Exclusion Monitor; EU Laeken; German wellbeing indicators; UNDP social exclusion indicators (2012); Bristol social exclusion matrix; Progress report on poverty, inequality, and social exclusion (Quebec); Canada's Community Well-being Index; Multidimensional Inequality Framework; Individual Deprivation Measure (International Women's Development Agency); Italy's survey on the inclusion of LGBT people in accessing work and in the workplace. |
| **Education (educational attainment)** | Social Exclusion Monitor; EU Laeken; German well-being indicators; UNDP social exclusion indicators (2012); Bristol social exclusion matrix; Progress report on poverty, inequality, and social exclusion (Quebec); Canada's Community Well-being Index; Global Multidimensional Poverty Index; Multidimensional Inequality Framework; Individual Deprivation Measure (International Women's Development Agency). |
| **Health (disability status of individuals and members of their household)** | Social Exclusion Monitor; EU Laeken; German well-being indicators; UNDP social exclusion framework (2012); Progress report on poverty, inequality, and social exclusion (Quebec); Türkiye National Indicators for the Rights of Persons with Disabilities (under the Convention of the Rights of Persons with Disabilities); Global Multidimensional Poverty Index; Multidimensional Inequality Framework; Individual Deprivation Measure (International Women's Development Agency); Mexico's National Survey on the Dynamics of Household Relationships (ENDIREH). |
| **Access to infrastructure** | UNDP social exclusion framework (2012); UNDP social inclusion framework (Bosnia and Herzegovina); Netherlands social exclusion framework; Bristol social exclusion matrix; Progress report on poverty, inequality and social exclusion (Quebec); Multidimensional Inequality Framework; Individual Deprivation Measure (International Women's Development Agency); Mexico's National Survey on the Dynamics of Household Relationships (ENDIREH). |

*Source: World Bank staff compilation from UNECE (2022).*



**Social exclusion can be measured through a collection of separate indicators, or through a composite index.** Individual indicators are user-friendly and straightforward to construct, but the co-existence of several indicators and the lack of a cohesive overall assessment may be problematic. Indices, on the other hand, synthesize multiple indicators into a single figure. This offers a clear summary, but requires more technical work to construct the index, and can potentially obscure certain details. Indices require careful weighting and standardization of indicators, to ensure a meaningful representation of relevant domains of social exclusion. They must accurately reflect improvements or deteriorations in their underlying components, and it must be possible to disaggregate them, so as to assess the impact of each indicator on the overall index. Moreover, to guide policy effectively, indices should enable spatial comparisons and the tracking of changes over time.

## 5.2. MEASURING SOCIAL EXCLUSION IN THE EU

**In Europe, the main concept of social exclusion is based on the at-risk-of-poverty or social exclusion (AROPE) indicator described in section 2.1.4.** Like the AROP rate, the AROPE indicator is derived from EU-SILC survey data. The survey collects information on severe material and social deprivation at the individual and household levels, by determining the ability to afford a set of 13 items listed in Table 2. An individual is considered severely materially and socially deprived if they, or their household, cannot afford at least seven such items. Concurrently, low work intensity at the household level is calculated from survey data on the age of household members, their self-reported economic status, their income from various sources, and their usual working hours during the months in which they were employed. The combination of severe deprivation and low work intensity is reflected in the overall AROPE rate.

**The AROPE indicator is mainly representative at the NUTS 2 level, which hinders the development of policy interventions targeted at smaller areas.** Its lack of representativeness at smaller geographical scales may obscure pockets of social exclusion within a region, and limits its usefulness for the preparation and execution of strategies to address regional challenges.

**Most EU countries use the AROPE indicator as a core measure of social exclusion, supplementing it with a variety of locally relevant indicators.** This approach makes it possible to both capture local conditions and conduct cross-country comparisons. For example, Germany starts from the AROPE indicator, and adds other measures related to housing, health, well-being, and quality of life, as well as indicators specific to children. Well-being is assessed through 46 indicators across 12 dimensions; no hierarchical weighting is applied, and each indicator is treated as equally significant. Well-being metrics are then displayed via an interactive dashboard, employing various forms of data visualization (UNECE, 2022).

**Moreover, EU member states have devised a range of methodologies to explore social exclusion beyond Eurostat's definitions and surveys, and to delve into spatial disparities at finer geographical levels.** To this end, data collection efforts extend beyond the NSIs and involve central and local governments, including line ministries responsible for education, labor, and social policy. Relevant sources of data encompass the national accounts, the population census, and various administrative databases. Moreover, NSIs have been increasingly collaborating with academic institutions, developing sophisticated statistical approaches.

**As explored in section 5.1, the concept of social exclusion is complex, leading to the use of a wide spectrum of indicators across the EU.** Unlike poverty measurements, which primarily focus on monetary income, those pertaining to social exclusion must account for various aspects of life, such as

# 5. REGISTER-BASED MEASUREMENTS OF SOCIAL EXCLUSION IN CROATIA

health, education, housing, and demographics (UNECE, 2022). Additionally, social exclusion is dynamic and context-specific, which further complicates the development of a unified approach to measuring it—for example, the factors contributing to social exclusion in a rural community may differ from those at play in a metropolitan area. Table 15 lists domains commonly considered when measuring social exclusion in the EU, and examples of indicators within each domain.

**Table 15.** Example of social exclusion domains and related indicators

| Domain | Indicators | EU member state(s) |
|---|---|---|
| **Income** | Disposable income | Belgium, Denmark, Germany, Spain, Portugal, Sweden |
| | Fiscal revenue per capita | Estonia, Poland |
| | Average salary (hourly, monthly, gross, net) | France, Poland, Slovenia |
| | GDP per capita | Germany, Spain |
| **Demographics** | Population density | Most member states |
| | Population distribution by age groups | France, Germany, Greece, Slovenia |
| | Age dependency ratio | Hungary, Luxembourg, Slovenia |
| | Ethnicity | Romania, Slovakia |
| **Health** | Number of hospital beds per 1,000 inhabitants | Czechia, Italy, Poland, Romania |
| | Number of physicians and nurses per 1,000 inhabitants | Czechia, Italy, Poland, Romania, Spain |
| | Mortality rate (infant, underfive) | Czechia, Italy, Poland, Romania, Spain |
| **Education** | Distribution of population by education attained | Czechia, France, Italy, Latvia, Romania, Sweden |
| | Student-teacher ratio | Italy, Poland |
| | PISA score | Czechia, Italy, Poland |
| **Housing and energy** | Price per square meter for rental housing | Denmark |
| | Overcrowding rate | Portugal |
| | Number of dwellings per 1,000 inhabitants | Slovenia |
| | Availability of heating in the house | Czechia, Italy, Romania, Spain |
| **Labor market** | Unemployment rate | Czechia, Denmark, Hungary, Italy, Poland |
| | Labor force participation rate | Czechia, Italy, Poland, Romania, Spain |
| | Activity rate | Austria |
| | Job density | Germany |

*Source: World Bank compilation.*

**In addition, many EU member states have developed multidimensional indices of poverty and social exclusion, adopting a variety of approaches.** A major challenge to building such an index is the lack of consensus on the indicators that it should encompass. Stakeholders—from central to local governments and line ministries—may have different priorities and perspectives as to what dimensions are most relevant to well-being or development. Another challenge lies in weighing the domains and indicators within the index; their relative importance can vary depending on the local context and the objective of the index. Determining a set of weights may also involve a subjective judgment—e.g., about the value of education relative to health, or of employment relative to housing conditions. Table 16 offers a glimpse into the wide range of multidimensional indexes used across the EU.

**Table 16.** Examples of multidimensional indexes in the EU

| Index | EU member state(s) |
|---|---|
| Living Condition Index | Austria |
| Wealth index based on fiscal revenue | Belgium |
| Poverty Index | Belgium |
| Index of multi-deprivation | Bulgaria, Croatia, Portugal (Lisbon) |
| Vulnerability to job loss index | France |
| Deprivation index | France, Ireland |
| Atlas of gender inequality | Germany |
| Regional atlas of income, human capital, social exclusion | Germany |
| Composite index of wellbeing | Greece |
| Segregation index | Hungary |
| Social and material deprivation index | Italy |
| Socio-economic index | Luxembourg |
| Local human capital index | Poland, Romania |
| Regional development index | Portugal |
| Atlas of marginalized communities | Romania |
| Atlas of Roma communities | Slovakia |
| Demographic vulnerability index | Sweden |

*Source: World Bank compilation.*

## 5.3. EXISTING APPROACH TO MEASURING SOCIAL EXCLUSION IN CROATIA

**The AROPE rate is the primary measure of social exclusion in Croatia.** In 2023, the national AROPE rate stood at 20.7 percent,[79] with significant variations across NUTS 2 regions: 31.3 percent in Pannonian Croatia, 11.9 in the City of Zagreb, 18.9 percent in Adriatic Croatia, and 18.6 percent in Northern Croatia.[80] These figures highlight disparities even within the broader NUTS 2 regions.

79 https://ec.europa.eu/eurostat/databrowser/view/ilc_peps01n/default/table?lang=en&category=livcon.ilc.ilc_pe.ilc_peps

80 https://ec.europa.eu/eurostat/databrowser/view/tgs00107/default/table?lang=en&category=t_ilc.t_ilc_pe

# 5. REGISTER-BASED MEASUREMENTS OF SOCIAL EXCLUSION IN CROATIA

**In addition, Croatia has adopted further methods to assess social exclusion at the municipal (LAU) level, such as the Development Index (DI) and the Index of Multiple Deprivation (IMD).** The MRDEUF introduced the DI in 2010 to periodically assess the level of development in local and regional government units.[81] This composite indicator is calculated from metrics such as unemployment rate, income per capita, local government revenues per capita, general population movements, tertiary education rate, and the aging index.[82] As per Article 36 of the Law on Regional Development of the Republic of Croatia, municipalities below the average development level, as indicated by the DI, are classified as "assisted areas." Since its introduction, the DI has been updated in 2013, 2017, and 2024, and informed four government decisions on the classification of local and regional self-government units based on their development status. Moreover, the MRDEUF developed an IMD based on indicators related to demography, social status, educational status, economic strength and potential of local units, population density, and a coefficient for war-affected areas.[83] Such index was used in 2015 to identify the country's most deprived areas, which then received financial resources through the EU-funded Program of Integrated Physical, Economic, and Social Regeneration of Small Cities in War-Affected Areas. In 2017, the MRDEUF, in collaboration with the WB, strengthened the IMD to enable an assessment of deprivation in terms of economic factors, social factors, and access to services across eight subdomains—namely, labor market, fiscal capacity, economic development, social protection, health and education, demography, social services, and physical infrastructure. The construction of the IMD relied on data from the 2011 Population Census, national statistics, and selected administrative sources.[84] However, since the IMD was produced on a one-off basis, its potential for use in continuous policy monitoring is limited.

## 5.4. RECOMMENDATIONS FOR CROATIA

**The recommendations in this section were informed by detailed consultations with the MLPSFSP, in addition to desk research and data analysis.** These efforts aimed to tailor the recommendations to the needs of the MLPSFSP, and to ensure that they can be both applicable and achievable.

### 5.4.1. Simplified AROPE rate

**As discussed in section 2.1.4, producing an AROPE rate fully compliant with Eurostat's definition requires administrative data on three components: (i) the AROP rate, (ii) individuals "severely materially and socially deprived", and (iii) households with "very low work intensity."** Croatia's administrative databases have the potential to provide data on components (i) and (iii), but not on component (ii)—severe material and social deprivation.

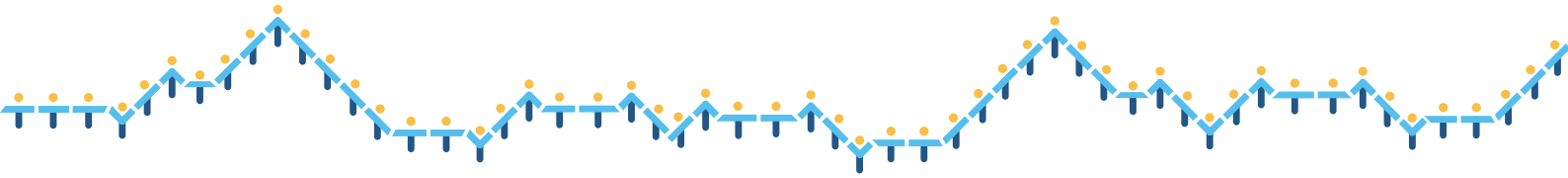
**Based on the official definition of the indicator, establishing who is "severely materially and socially deprived" largely relies on self-reported survey data.** Table 2 in section 2.1.2 details the 13 so-called deprivation items relevant to "severe material and social deprivation." In Croatia, administrative databases lack information on several of them, such as the capacity to "get together with friends and family for a drink/meal at least once a month", the ability to "replace worn-out clothes with new ones", and possession

81 In accordance with the Law on Regional Development of the Republic of Croatia (Official Gazette, nos. 147/14, 123/17, and 118/18)

82 https://razvoj.gov.hr/o-ministarstvu/regionalni-razvoj/indeks-razvijenosti/112

83 https://razvoj.gov.hr/UserDocsImages/O%20ministarstvu/Regionalni%20razvoj/Odluka%20o%20pilot%20podru%C4%8Djima.pdf

84 https://razvoj.gov.hr/UserDocsImages/O%20ministarstvu/Regionalni%20razvoj/Index%20of%20Multiple%20Deprivation%20-%20Conceptual%20framework_22_10_2018.pdf



of "two pairs of properly fitting shoes (including a pair of all-weather shoes)." Therefore, Eurostat's indicator of "severe material and social deprivation" cannot be fully replicated using administrative data alone.

**On the other hand, Population Register and Tax Income data enable the calculation of AROP rates closely aligned with Eurostat's definition.** The relevant methodology, detailed in Chapter 4, accounts for all income components specified by Eurostat, except for private transfers among households. Although Tax Income data has some gaps, particularly about expenditures by self-employed individuals and accommodation renters subject to the lump-sum tax regime, the missing data can be imputed using statistical models commonly employed in the EU.

**Measures of "very low work intensity" could be estimated using data from the Population Register and the Tax Income database.** The relevant population for this estimate includes individuals aged 18-64, except four specific groups: students aged 18-24, retirees, pension income recipients,[85] and those aged over 60 who do not work and live in households that rely primarily on pension income.[86] The Draft Law of the Central Register of the Population suggests that the data in the Population Register (refer to Annex A2) should be sufficient for identifying such groups (Table 17).

**Table 17.** Administrative data sources for identifying individuals living in households with "very low work intensity"

| Population group | Data source |
|---|---|
| Individuals aged 18-64 | Population Register |
| Students aged between 18-24<br>• Secondary level<br>• Tertiary level | Population Register |
| Retirees | Population Register |
| Pension income recipients | Population Register and Tax Income (for monetary receipts from pensions) |
| People over 60 who do not work and live in households that rely primarily on pension income | Population Register and Tax Income (for employment status and monetary receipts from pensions) |
| For employees, the number of hours worked | Tax Income |
| For self-employed individuals and workers in "other activities" (in Croatian: druga djelatnost), the number of months worked | Tax income (note: this source does not include information on number of hours worked for these groups of workers. Therefore, it is assumed that in the months in which they were active, they worked on a full-time basis) |
| Household identification | Population Register |

*Source: World Bank staff elaboration.*

**The proposed AROPE-Simplified indicator incorporates two of the three components of Eurostat's AROPE concept: (i) the AROP rate, and (iii) households with "very low work intensity."** This streamlined version excludes the "severely materially and socially deprived" element of the full AROPE measure.

---

**85** Excluding survivors' pensions, disability pensions, and pensions from individual plans.

**86** Excluding survivors' pensions, disability pensions, and pensions from individual plans.

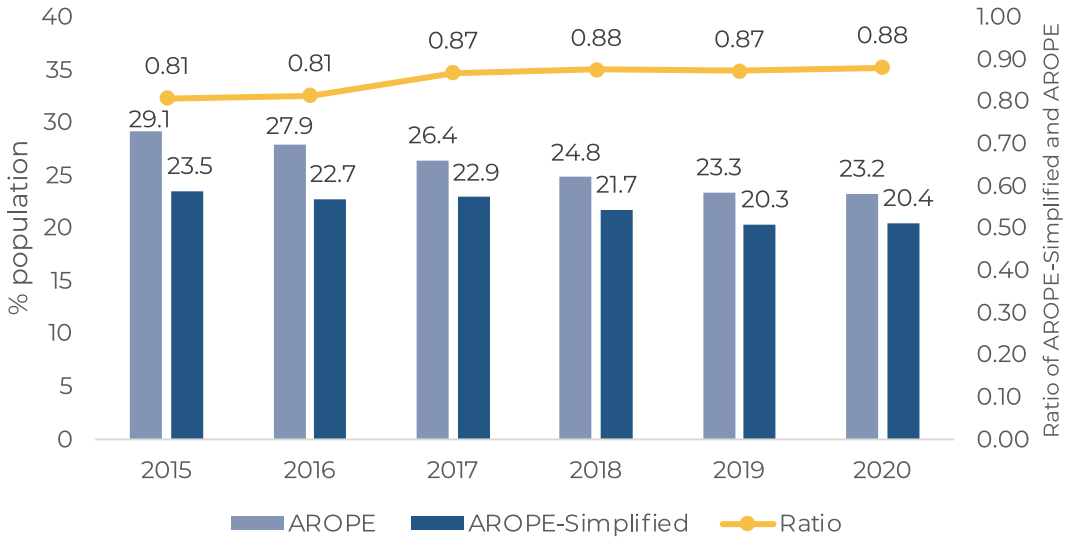
# 5. REGISTER-BASED MEASUREMENTS OF SOCIAL EXCLUSION IN CROATIA

However, in theory, whether households can afford the 13 deprivation items in Table 2 shows whether they can generally afford certain goods and services—such as furniture, clothes, shoes, and holidays—in a way that is arguably correlated with their disposable income. In addition, since labor income tends to account for a significant portion of total income, the capacity to afford such items is likely to have a strong link with a household's work intensity. Therefore, the AROPE-Simplified indicator is expected to be highly correlated with the AROPE indicator.

**A comparison of the AROPE and AROPE-Simplified rates calculated from EU-SILC survey data reveals small differences, which warrant further examination.** As Eurostat revised the definition of "very low work intensity" in 2021,[87] we compared the AROPE-Simplified and AROPE rates for Croatia using both the old and new definitions (Figure 9). The results suggest that the AROPE-Simplified rates closely mirror the AROPE rates under Eurostat's revised definition of "very low work intensity." Panel 9a displays the comparison for 2015–2020 under the old definition, with the ratio between the AROPE-Simplified and AROPE rates fluctuating between 0.81 and 0.88. Panel 9b shows the comparison after retroactively applying the updated definition, and reveals a significant convergence of the two rates over time—from 0.90 in 2015 to 0.97 in 2023. For example, in 2023, the AROPE-Simplified rate stood at 20.1 percent, only 0.6 percentage points below the AROPE rate. It is important to further test the AROPE-Simplified concept using more recent EU-SILC survey data, as it becomes available; and to calculate the two rates for various population groups (e.g., using breakdowns by gender and age), to determine whether the small differences observed in aggregate also occur across subgroups.

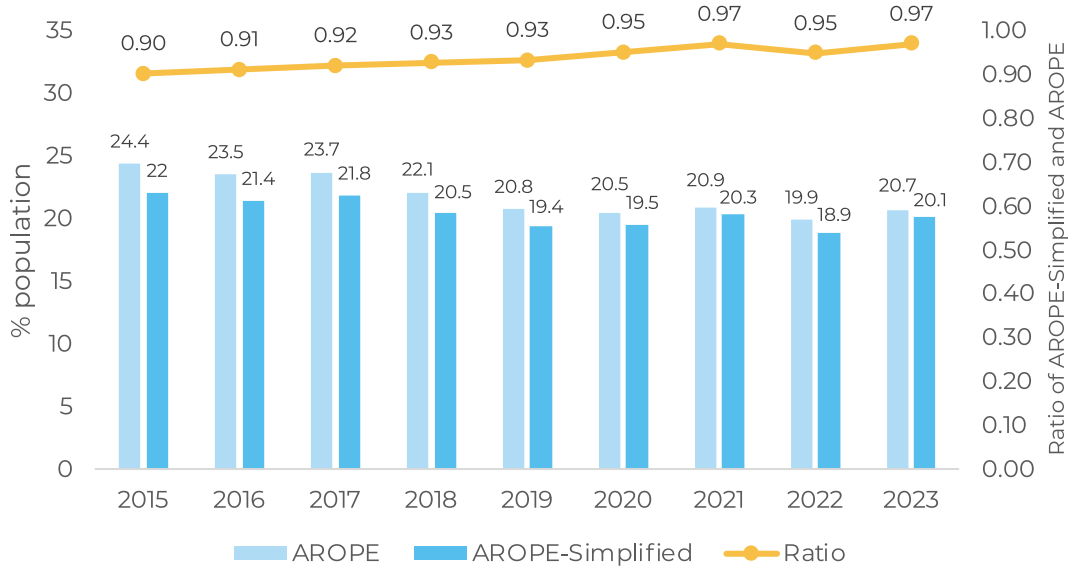
**Figure 9.** Comparison of AROPE and AROPE-Simplified rates in Croatia based on EU-SILC data

Panel 9a. Comparison under Eurostat's old definition of "very low work intensity"




87 Until 2021, working-age adults were defined as those aged 18 to 59, excluding students aged 18 to 24. The revised definition of working-age population includes people up to the age of 64, excluding inactive people aged 60 to 64 living in a household where pensions are the main source of income. It is unclear if this change contributes to the narrowing gap between the AROPE-Simplified and AROPE rates over time.

Panel 9b. Comparison under Eurostat's revised definition of "very low work intensity"




*Source: World Bank staff calculation based on Eurostat: https://ec.europa.eu/eurostat/databrowser/view/ilc_pees01n__custom_11602706/default/table?lang=en*

## 5.4.2. Social exclusion indicators by domain

**The AROPE-Simplified concept may not fully capture the complex nature of social exclusion, suggesting a need for a broader range of domains and indicators.** The AROPE-Simplified concept focuses solely on disposable income and low work intensity. Even the AROPE concept, which considers a broader range of deprivation items, overlooks several key aspects of social exclusion—such as those related to health, education, and access to basic services. A more comprehensive approach would have to consider multiple domains and indicators, to support the design of sector-specific policies and investments in line with the government's objectives.

**In consultation with the MLPSFSP, nine suitable domains have been identified, along with related indicators that can be derived from existing administrative data sources.** Table 18 presents the domains—i.e., household income, social protection, labor market, education, health, demography, infrastructure, business activity, and a "miscellaneous" category—along with the suggested indicators and their corresponding data sources. Indicators within each domain were selected based on their policy relevance (i.e., how accurately they reflect the impact of policy interventions) and feasibility (i.e., how practical it is to estimate them based on the available data).

**Considering the technical complexity of calculating certain indicators, it is recommended that the MLPSFSP focus its time and financial resources on producing those most relevant to its needs and priorities.** Estimating certain indicators, such as the number of primary-care doctors per 1,000 inhabitants at the municipal level, is relatively straightforward. Calculating others, however, entails dealing with complexities around definitions and the involvement of multiple stakeholders. For example, one proposed indicator concerns the density of buses operated in a municipality, as a proxy for the accessibility of public transport. In Croatia, however, the definition of public transport might have to include buses run by private companies under local government contracts. Moreover, since bus routes often span several

municipalities, it is crucial to define the indicator accurately to prevent double counting. Finally, collecting data for this indicator and updating it annually would require coordination among various stakeholders, including local governments and private companies in each municipality, further complicating the task.

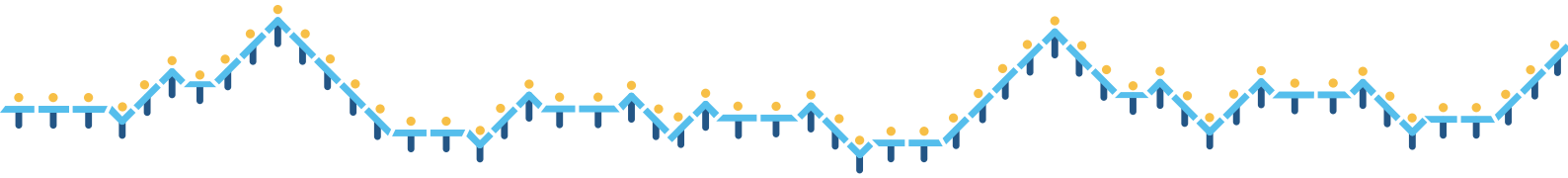
**Many of the proposed indicators can be constructed for particularly vulnerable population groups, including women, children, the elderly, and disabled people.** Group-specific indicators can be calculated from individual-level microdata, when necessary demographic information (such as gender, age, and disability status) is available. Specifically, all indicators based on Population Register and Tax Income data can be disaggregated by these characteristics.

**Certain domains and indicators can be linked to the objectives of several Action Plans for which the MLPSFSP is responsible:** the Action Plan for the Fight Against Poverty and Social Exclusion, the Action Plan for the Equalization of Opportunities for People with Disabilities, the Action Plan for the Implementation of the National Plan for Work, Protection at Work, and Employment, and the Action Plan for the Implementation of the National Plan for Gender Equality. Such plans include specific objectives such as alleviating poverty and social exclusion among vulnerable groups, preventing child poverty, improving labor market access for the unemployed and inactive, enhancing the employment system for people with disabilities, and improving the labor market status of women. Progress towards these goals can be captured by proposed indicators in the household income, social protection, and labor market domains.

**The proposed indicators can also serve as the basis for a composite index to measure social exclusion, if needed.** A detailed explanation of the process for developing such an index can be found in Annex A6. The approach is similar to that applied for creating the IMD, discussed in section 5.3, although it features domains and indicators more relevant to the needs of the MLPSFSP.

**Table 18.** Proposed indicators of social exclusion at the NUTS 3 and LAU levels

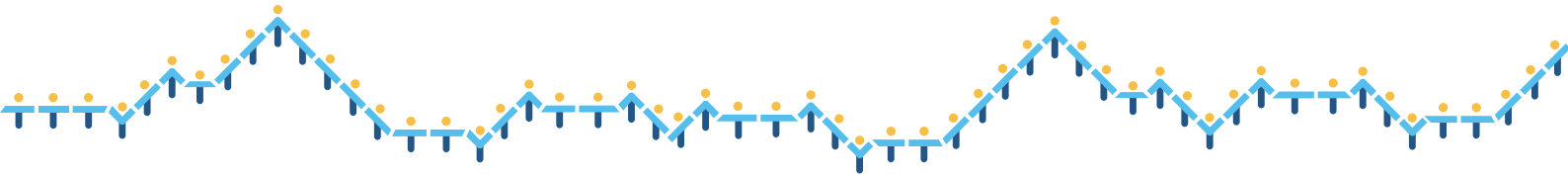
| Indicator name | Description | Data sources |
|---|---|---|
| **Domain: Household income** | | |
| Average disposable income | Mean household disposable income per adult-equivalent. | Population Register linked to Tax Administration data |
| Median disposable income | Median household disposable income per adult-equivalent. | Population Register linked to Tax Administration data |
| Average disposable income for the poorest 40% | Mean household disposable income per adult-equivalent among the poorest 40% of the population, defined as those whose household disposable income falls below the 40th percentile of the municipality/ town distribution | Population Register linked to Tax Administration data |
| Average disposable income for the poorest 40% before social transfers | Mean household disposable income before social transfers per-adult equivalent among the poorest 40% of the population, defined as above. | Population Register linked to Tax Administration data |
| Population without market income or pensions | Persons living in households that do not earn a market-based income or old-age pension, as a share of the population. | Population Register linked to Tax Administration data |



| | | |
|---|---|---|
| **Domain: Social protection** | | |
| Population receiving social transfers | Persons living in households that receive social transfers, as a share of the population. Social transfers include all means-tested transfers from the central or local government. | Population Register linked to Tax Administration data |
| Population living off social transfers mostly | Persons living in households where social transfers make up more than 50% of disposable income, as a share of the population. Social transfers include all transfers from the central or local government. | Population Register linked to Tax Administration data |
| Population with social transfers from Law on Social Welfare and Law on Supplement for Inclusion | Persons living in households that receive social transfers as regulated by the Law on Social Welfare and the Law on Supplement for Inclusion, as a share of the population. | Population Register linked to Tax Administration data |
| **Domain: Labor market** | | |
| Employment rate | Employment rate for persons aged 15-64. Employed persons are those who pay mandatory pension contributions in the relevant year, in any amount. | Population Register linked to Tax Administration data |
| Unemployment rate | Unemployment rate for persons aged 15-64. | Population Register and the Croatian Employment Service, for the number of people participating in the labor market; Croatian Employment Service, for the number of unemployed persons. |
| Long-term unemployment rate | Long-term unemployment rate for persons aged 15-64. Long-term unemployed persons are those who have been unemployed for a year or longer. | Population Register, for the population aged 15-64; Croatian Employment Service, for the number of long-term unemployed aged 15-64 |
| Participation rate | Participation rate for persons aged 15-64. Persons who participate in the labor market are those who are either employed or unemployed. | Population Register, for the population aged 15-64; Population Register linked to Tax Administration data and Croatian Employment Service, for the number of labor market participants aged 15-64 |
| **Domain: Education** | | |
| Kindergarten enrollment | Kindergarten enrollment rate (%) for children aged 1-5. | Population Register for the population aged 1-5; Ministry of Science and Education for the number of children aged 1-5 enrolled |
| Matura exam performance | Failing rate (%) at Matura exam. | National Center for Evaluation of Education |
| Teachers per pupil | Number of teachers per pupil in elementary school. | Ministry of Science and Education |

# 5. REGISTER-BASED MEASUREMENTS OF SOCIAL EXCLUSION IN CROATIA

| Domain: Health | | |
|---|---|---|
| Density of medical doctors | Number of medical doctors per 1,000 inhabitants. | Population Register, for population size; Croatian Health Insurance Fund, for the number of medical doctors |
| Density of family medicine specialists | Number of family medicine specialists per 1.000 inhabitants. | Population Register, the population size; Croatian Health Insurance Fund and Croatian Medical Chamber (Digital Atlas of Croatian Medicine) for the number of specialists |
| Density of pediatricians | Number of pediatricians per 1,000 children. | Population Register, for population size of children; Croatian Health Insurance Fund and Croatian Medical Chamber (Digital Atlas of Croatian Medicine) for the number of specialists |
| Density of school and adolescent medicine specialists | Number of school and adolescent medicine specialists per 1,000 children and adolescents. | Population Register, for size of the population of children and adolescents; Croatian Health Insurance Fund and Croatian Medical Chamber (Digital Atlas of Croatian Medicine) for the number of specialists |
| Density of gynecologists | Number of gynecologists per 1,000 women. | Population Register, for female population size; Croatian Health Insurance Fund and Croatian Medical Chamber (Digital Atlas of Croatian Medicine) for the number of specialists |
| Density of psychiatrists | Number of psychiatrists per 1,000 inhabitants | Population Register, for population size; Croatian Health Insurance Fund and Croatian Medical Chamber (Digital Atlas of Croatian Medicine) for the number of specialists |
| Density of child and adolescent psychiatrists | Number of child and adolescent psychiatrists per 1,000 children and adolescents | Population Register, for size of the population of children and adolescents; Croatian Health Insurance Fund and Croatian Medical Chamber (Digital Atlas of Croatian Medicine) for the number of specialists |
| Density of geriatricians | Number of geriatricians per 1,000 inhabitants older than 75 | Population Register, for the size of the population older than 75; Croatian Health Insurance Fund and Croatian Medical Chamber (Digital Atlas of Croatian Medicine) for the number of specialists |
| Density of dentists | Number of dentists per 1,000 inhabitants. | Population Register, for population size; Croatian Health Insurance Fund, for the number of dentists |



| | | |
|---|---|---|
| Nearest primary health center | Distance (in meters) to the nearest primary health center—i.e., the distance between the centroid of the municipality/town and the health center (per coordinates of its address), based on Geographical Information System (GIS) data. | Croatian Health Insurance Fund, for addresses of primary health centers; State Geodetic Administration, for GIS data |
| Nearest hospital | Distance (in meters) to the nearest hospital—i.e., distance between the centroid of the municipality/town and the hospital (per the coordinates of its address), based on GIS data. | Croatian Health Insurance Fund, for addresses of hospitals; State Geodetic Administration, for GIS data |
| **Domain: Demography** | | |
| Population change | Change in population (i.e., number of residents), in %, from year T-1 to year T. | Population Register |
| Working-age population change | Change in working-age population (aged 15-64) from year T-1 to year T. | Population Register |
| Dependency ratio | Number of persons younger than 15 or older than 64 per one working-age person (aged 15-64). | Population Register |
| Population density | Number of inhabitants per square kilometer. | Population Register, for population size; Croatian Bureau of Statistics, for surface area |
| **Domain: Infrastructure** | | |
| Paved road density | Length of paved roads (in kilometers, based on GIS data) per square kilometer. | Ministry of Maritime Affairs, Transport and Infrastructure, for the length of paved roads; Croatian Bureau of Statistics, for surface area |
| Access to bus transportation (*,**) | Number of bus lines stopping in a municipality or town per 1,000 inhabitants | Local governments and transport companies, for number of bus lines (potential source: Ministry of Maritime Affairs, Transport and Infrastructure, registry of permits from the Law on road transport); Population Register, for population |
| Water pipe density | Length of water pipes (in kilometers) per square kilometer of a municipality/town's area. | Croatian Waters, for the length of water pipes; Croatian Bureau of Statistics, for surface area |
| Sewage coverage (*) | Share of households connected to sewage | Local governments and companies responsible for sewage systems, for sewage connection status; Population Register, for population |

# 5. REGISTER-BASED MEASUREMENTS OF SOCIAL EXCLUSION IN CROATIA

| | | |
|---|---|---|
| **Domain: Business activity** | | |
| Density of businesses | Number of active business entities per inhabitant. | Population Register, for population size; Croatian Bureau of Statistics, for the number of active business entities |
| Density of crafts | Number of active crafts per inhabitant. | Population Register, for population size; Croatian Bureau of Statistics, for the number of active crafts |
| **Domain: Miscellaneous** | | |
| Personal cars | Number of registered personal cars per inhabitant. | Population Register, for population size; Ministry of Interior, for the number of registered personal cars |
| Persons subject to foreclosure | Persons subject to forced collection of unpaid debt, as a share of the population. | Population Register, for population size; Financial Agency (FINA), for the number of persons subject to debt enforcement |

*Source: World Bank staff elaboration.*

*Note: (*) MLPSFSP should assess the feasibility of collecting this data.*

*(**) The definitions and concepts of bus lines and public transport should be aligned with the Law on road transport.*[88]

88 Law on road transport (Official Gazette no. 41/18, 98/19, 30/21, 89/21, 114/22)

### 5.4.3. Topic-specific surveys

**Although administrative databases offer numerous benefits, they alone may not capture all aspects of social exclusion or offer an accurate snapshot of vulnerable populations.** Administrative data sources contain extensive and varied socio-economic information, ranging from human capital to employment and income. However, they have limitations, particularly when it comes to capturing the complexities of social exclusion and the experiences of vulnerable groups. Notably, such databases often rely on records that may not cover certain marginalized populations, or may not delve into issues that disproportionately affect them. For instance, incidents of sexual harassment or violence against girls and women are typically underreported in official records due to stigma, fear of retribution, or lack of trust in the authorities. Similarly, transient or marginalized groups, such as the Roma or the homeless, may not be consistently tracked in administrative systems, leading to a lack of reliable data on their circumstances.

**Thus, surveys are a critical complement to administrative data, to address its shortcomings and fill the gaps.** Through carefully designed questionnaires and interviews, surveys can uncover the prevalence of issues, such as violence and abuse, that are not adequately captured by other data sources. They can also reach population groups that are often invisible in administrative databases, providing insights into their needs, challenges, and the barriers they face. Appropriate methodologies, such as anonymous responses or community-based participatory research, can help create a safe space for participants to share their experiences without fear of stigma or reprisal. The result can be a more accurate and comprehensive picture of social issues, informing more effective and targeted policy interventions.

**When planning topic-specific surveys, the MLPSFSP should carefully consider the key drivers of their costs: i) level of representativeness, ii) level of detail of the questionnaire, and iii) methods and frequency of data collection.** With regard to i), for instance, the previous report in this series indicated that in Croatia, a representative sample at the NUTS 1 (national) and NUTS 2 (regional) levels might require approximately 5,000 households and 12,500 households respectively. However, for representativeness at the NUTS 3 (county) level, the sample size should range between 23,000 and 28,000 households. The number of households necessary for each survey depends on their objectives and quality requirements, which in turn impact the related costs. As a reference, the CBS recently projected that carrying out an EU-SILC survey at the NUTS 3 level would cost between €750,000 and €1 million.

**The comprehensiveness of the questionnaire is another critical aspect that affects the cost of conducting surveys.** A broad and in-depth questionnaire requires a significant investment of time from both interviewers and respondents, potentially over several hours and multiple visits. On the other hand, a survey consisting of a 15-minute interview would be cheaper, but might not provide equally comprehensive data. The depth and breadth of the questionnaire must be balanced against the financial constraints and goals of the survey, to ensure that it is both economically viable and effective in gathering the necessary information.

**Moreover, the choice of data collection methods has implications on both the cost of the survey and the population groups covered.** In-person interviews, while offering a personal touch and potentially higher response rates, are usually expensive due to travel and time-related costs. Telephone surveys and online forms can be cheaper and logistically simpler, but might not be as effective at reaching certain population groups, such as those who lack reliable internet access or prefer face-to-face communication. Each method has its own advantages and disadvantages (Box 13), which the MLPSFSP must weigh against the survey's objectives and the characteristics of the target population to determine the most cost-effective and representative approach.

# 5. REGISTER-BASED MEASUREMENTS OF SOCIAL EXCLUSION IN CROATIA

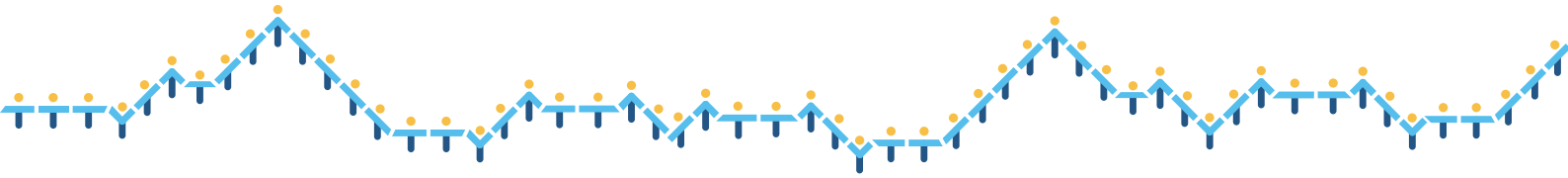
**Box 13.** Advantages and disadvantages of various methods to collect survey data

**Choosing the right method of data collection is crucial to the success of a survey.** The technique adopted must ensure extensive coverage, a high response rate, and accurate data gathering, while minimizing the burden on respondents and keeping costs manageable. Since achieving all such goals at once is not possible, the selection often depends on the specific requirements of the survey.

**In-person interviews enable the interviewer to increase the response rate and improve the quality of the information collected, by means of personalizing the interview and interpreting questions and survey concepts** (United Nations, 2022). This method allows for detailed data collection through direct interaction, which can lead to higher-quality data and fewer misunderstandings. It is particularly useful for complex or sensitive topics on which the interviewer's assistance is beneficial, or when respondents have a low level of literacy. On the other hand, such surveys are often expensive and time-consuming, due to the need for trained personnel and travel. They are also vulnerable to social desirability bias, whereby respondents may offer an answer that they deem socially acceptable, rather than one that is true to their beliefs.

**Phone surveys are accessible and convenient for respondents, and can increase response rates.** The computer-assisted telephone interview (CATI) is a technique through which an interviewer asks questions during a phone conversation (over a fixed or mobile line), and enters the responses into a computer system. This method provides good response rates at a moderate cost. Telephone interviews are quicker and less expensive than face-to-face interviews, as they eliminate travel expenses and reduce the time required to reach respondents. They also facilitate quality control, as the interview process can be monitored effectively. However, they entail constraints to interview length and questionnaire complexity, as respondents tend to have less tolerance for long or intricate interviews when they are conducted over the phone rather than in person. The use of mobile phones has also had an impact on telephone interviews: developing a sample frame that adequately covers the target population has become more complex; and response rates have dropped slightly, as caller ID functions make mobile phone users less likely to answer calls from unknown numbers than landline users (although pre-survey communication and awareness campaigns, or prior notice to interviewees about the upcoming phone call, may mitigate this issue). In this context, CATI remains an effective method for conducting recurring surveys and follow-ups, especially when respondents have previously supplied reliable contact details and agreed to participate.

**Online surveys are cost-effective and can reach a wide audience quickly** (Evans and Mathur, 2005). Online surveys enable easy access to a diverse population across different regions or countries at a very low cost. They also allow for anonymity, which can reduce social desirability bias (Duffy et al. 2005). However, online surveys may prompt short and simplified responses, and their quality can suffer due to multiple submissions, non-serious responses, and lack of physical oversight (Chang and Vowles 2013). They also entail a risk of excluding individuals who do not have internet access, causing coverage errors and bias in the results (Andrade 2020); and elicit weaker engagement from respondents relative to other methods, resulting in lower-quality data (Evans and Mathur 2018, Mavletova 2013). Therefore, online surveys are ideal for large-scale data collection efforts that aim for a vast geographical reach on a constrained budget.



**For reference, the 2021 Population Census in Croatia adopted a mixed approach.** The onset of the COVID-19 pandemic caused a six-month postponement of the reference date, and data was gathered through both online submissions and in-person interviews. First, the population was invited to submit information via the e-Citizens electronic system between September 13th and 26th, 2021. Then, census enumerators conducted face-to-face interviews from September 27th to November 14th. As in some other EU member states, the usual post-enumeration survey was not conducted, due to persistent epidemiological risk.[89]

89 https://zadarskilist.novilist.hr/novosti/hrvatska/osim-hrvatske-jos-24-clanice-eu-nisu-provele-kontrolni-popis/

# 6. INSTITUTIONAL SET-UP FOR TRACKING POVERTY AND SOCIAL EXCLUSION

**This chapter explores potential institutional arrangements for the estimation, analysis, and dissemination of poverty and social exclusion indicators at the NUTS 3 and LAU levels.** While the MLPSFSP has a crucial role in every scenario, several options envisage varying degrees of involvement from other key stakeholders, namely: the Tax Administration, the CBS, and the MJPADT.

## 6.1. ACTIVITIES AND REQUIREMENTS

**Regular tracking of poverty and social exclusion hinges on the execution of three key activities:**

1. Calculation of AROP and AROPE-Simplified rates: Such calculations are based on microdata from the Population Register and Tax Income databases, both managed by the Tax Administration.

2. Development of social exclusion indicators or index: This entails constructing a set of indicators or a composite index that accurately reflect the multifaceted nature of social exclusion, in a way that is relevant to the MLPSFSP. This effort can rely on administrative records from the Population Register, the Tax Income database, and other data sources such as data from the CBS, the Ministry of Science and Education, the Croatian Health Insurance Fund, and the Ministry of Maritime Affairs, Transport, and Infrastructure.

3. Production of visual materials for reporting: Such materials—e.g., tables, charts, maps, and infographics—must succinctly convey the key findings from the two steps above to a diverse audience, including policymakers, stakeholders, and the general public. Their goal is to ensure that insights gained from the data are accessible and comprehensive, and can feed into relevant National Plans while facilitating public awareness and informed decision-making on matters related to poverty and social exclusion.

**The institutional arrangements for executing such activities will be crucial.** The ultimate success of the efforts above will depend on inter-institutional agreements on the roles and responsibilities of each stakeholder. The appropriate institutional setup should satisfy three key requirements:

1. Data Accessibility: It is crucial to ensure reliable access to the necessary microdata, which serves as the basis for all subsequent activities.
2. Technical Capacity: The activities above require technical expertise among the stakeholders involved. In particular, the MLPSFSP, the Tax Administration, the CBS, and the MJPADT must have the capability of merging data, assessing its quality, and calculating indicators.
3. Collaboration: It will be vital for the MLPSFSP, the Tax Administration, the CBS, and the MJPADT to collaborate on activities that may extend beyond their traditional scope of work.

## 6.2. POTENTIAL INSTITUTIONAL SET-UPs

**This section presents four potential institutional configurations for the implementation of the three activities listed in section 6.1.** Such options were identified through discussions with the MLPSFSP and relevant stakeholders. The order in which they are discussed does not indicate priority or preference.

**Option 1: MLPSFSP-led implementation.** In this scenario, the MLPSFSP would be in charge of all three activities. This centralized approach requires significant investment by the MLPSFSP in staffing and data infrastructure.

**Option 2: Joint implementation by MLPSFSP and Tax Administration.** In this scenario, the Tax Administration would merge the data, calculate the AROP rate, and contribute to the estimation of certain social exclusion indicators, such as the AROPE-Simplified rate. The MLPSFSP would then calculate the remaining social exclusion indicators, and perform all reporting duties. This collaborative model aims to leverage the specialized expertise and resources of both ministries to enhance the management and analysis process.

**Option 3: Joint implementation by MLPSFSP and MJPADT.** This option harnesses the advanced infrastructure of the MJPADT with a view to improving data processing and analysis across all activities.

**Option 4. Joint implementation by MLPSFSP and CBS.** In this scenario, the involvement of the CBS would range from direct execution of key statistical activities to facilitation via the provision of access to data, expert advice, statistical methodologies, and other forms of support.

The next section provides more detail of each proposed option, along with an assessment of their advantages and disadvantages. Table 19 summarizes the roles of relevant stakeholders in each scenario.

**Table 19.** Stakeholders' roles in four potential institutional models

| Option for institutional set-up | Stakeholder responsible for: | | | | |
|---|---|---|---|---|---|
| | **Data access** | **Secure IT infrastruture** | **Estimates of AROP rate, AROPE-Simplified rate, income-based social exclusion indicators (technically more demanding)** | **Estimates of other social exclusion indicators (technically less demanding)** | **Monitoring system/ Data reporting** |
| **Option 1: MLPSFSP – led implementation** | MLPSFSP | MLPSFSP | MLPSFSP | MLPSFSP | MLPSFSP |
| **Option 2: Joint MLPSFSP – Tax Administration** | Tax Administration | Tax Administration | Tax Administration | MLPSFSP | MLPSFSP |
| **Option 3: Joint MLPSFSP – MJPADT** | MLPSFSP | MJPADT | MLPSFSP | MLPSFSP | MLPSFSP |
| **Option 4A: Joint MLPSFSP – CBS (as implementer)** | CBS | CBS | CBS | MLPSFSP | MLPSFSP |
| **Option 4B: Joint MLPSFSP – CBS (as facilitator)** | CBS | CBS | CBS's preapproved researcher(s) | MLPSFSP | MLPSFSP |

*Source: World Bank staff elaboration.*

# 6. INSTITUTIONAL SET-UP FOR TRACKING POVERTY AND SOCIAL EXCLUSION

## 6.2.1. MLPSFSP-led implementation

**To enable an approach under the sole responsibility of the MLPSFSP, the Ministry should meet certain critical requirements.** These concern, in particular, data access and technical capacity.

**Data Access Requirements:** It is recommended that the MLPSFSP obtain access to the Population Register and Tax Income microdata from the Tax Administration. To protect the privacy of individuals, both types of data can be anonymized in a consistent way, to enable their integration at the individual level for thorough analysis.

**Technical Capacity Requirements:** For data analysis purposes, it is recommended that the MLPSFSP establish:

- A robust technical team proficient in the use of advanced statistical software—such as Stata, SAS, R, or Python—to analyze data and develop, test, and validate econometric models.
- A strong monitoring and evaluation team, to collect, analyze, and track the trajectories of a wide range of poverty and social exclusion indicators.
- A secure IT environment, to prevent breaches and loss of sensitive data.

**The main advantage of an MLPSFSP-led approach lies in the centralized control of the Ministry over every aspect of the process.** Direct responsibility allows the MLPSFSP to set and adjust deadlines, prioritize tasks, and structure the process to be responsive to the Ministry's needs. This is particularly important when critical data is needed to inform the Ministry's policy decisions, allocate funding, or report to stakeholders, such as the MRDEUF.

**Key challenges for the MLPSFSP include access to the Population Register and Tax Income microdata, and staff recruitment.** As discussed in section 3.7, the draft Law on the Population Register is not explicit about the rights of government agencies to access the Register's microdata, as well as the procedure for doing so. It is important that the MLPSFSP follows up with the Tax Administration regarding this legal framework. Another challenge lies in the recruitment of staff with econometric modeling skills; despite recent wage increases in the public sector, attracting the right talent may still be difficult. It is recommended that the MLPSFSP develop comprehensive terms of reference for such roles, to ensure that they are intellectually stimulating, professionally rewarding, and entail a full use of their knowledge. Alternatively, the MLPSFSP may consider a collaboration with a research institute. It is equally important for the MLPSFSP to ensure that its monitoring and evaluation team works closely with the team overseeing the progress of the Ministry's National Plans, to streamline related activities and enhance efficiency.

## 6.2.2. Joint implementation by MLPSFSP and Tax Administration

**This collaborative strategy is designed to optimize the use of each agency's resources and expertise.** Since the Tax Administration manages the Population Register and the Tax Income database, in this scenario it would be responsible for estimating the AROP and AROPE-Simplified rates, as well as income-based social exclusion indicators, through statistical software and econometric modeling. Moreover, it would be responsible for maintaining a secure IT environment. On the other hand, the MLPSFSP would focus on calculating the remaining social exclusion indicators; interpreting and analyzing them; as well as reporting and disseminating the findings to the general public and to stakeholders, including the MRDEUF, in an accessible and informative manner. In this framework, the Tax Administration's data

processing capabilities ensure that the data is accurate and secure, while the MLPSFSP's strengths in analysis and communication make it possible for the findings to effectively inform policy and public understanding. This division of labor allows each ministry to work efficiently within its area of expertise, leading to more effective outcomes.

**For this arrangement to be effective, the Tax Administration should meet certain technical requirements:**

- Data analysis: The Tax Administration, or its designated agency APIS IT, should enhance the expertise of their data processing teams, ensuring proficiency in the use of advanced statistical software—such as Stata, R, or Python—to manage and analyze large data sets.
- Secure IT environment: The Tax Administration already has secure data infrastructure in place, with no need for additional investment in this area.
- Data sharing: The MoF should share the calculated indicators at the NUTS 3 and LAU levels with the MLPSFSP on an annual basis, ensuring the timely and consistent release of crucial data for further analysis and dissemination.

**For its part, the MLPSFSP would have to meet a different set of technical requirements:**

- Data collection and analysis: The MLPSFSP would have to estimate the social exclusion indicators that are not based on income. The relevant technical requirements depend on the chosen indicators or indexes. Most social exclusion indicators in Table 15 can be calculated with basic analytical tools, such as Microsoft Excel; however, indicators related to access to infrastructure may require GIS software skills. In addition, a composite index (if adopted) would require more complex data integration and analysis, and therefore a higher level of statistical expertise.
- Reporting: The MLPSFSP would be responsible for producing graphs, tables, and maps for reporting purposes. This requires proficiency in data visualization, potentially through the use of specialized software, to ensure that reports are informative and user-friendly.

**With this approach, the MLPSFSP would not need to access sensitive microdata directly, or to invest in secure data infrastructure.** Population Register and Tax Income data would be merged and processed within the secure IT system of the Tax Administration, with no need for further commitment of financial and human resources to ensure data protection. Therefore, the MLPSFSP could concentrate on other essential functions.

**The main uncertainty of this arrangement concerns the Tax Administration's commitment and readiness to undertake tasks that may fall outside its traditional purview.** The Tax Administration's dedication to enhancing its team's expertise is crucial, as data processing and econometric modeling require qualified staff. Personnel with the necessary foundational knowledge, and that could be trained for these tasks, may already work within the MoF— especially in the Bureau for Macroeconomic and Fiscal Analysis and Projections. Identifying and training them requires a deliberate allocation of staff time and budgetary resources. This proposed institutional setup also requires a formal inter-agency agreement between the MLPSFSP and MoF/Tax Administration, clearly defining responsibilities, budgets, and timelines for each party.

### 6.2.3. Joint implementation by MLPSFSP and MJPADT

**The institutional arrangement discussed in this section envisages using the MJPADT's digital infrastructure for data processing.** Per the draft Amendments to the Law on State Data Infrastructure,[90] government entities will be required to use a central interoperability system for data exchange across various registers. Such a system includes a protocol for data users to request authorization to access data, and for data holders to grant it. The data is then securely transmitted to other government agencies through an encrypted connection.

**The MJPADT's Data Lake digital infrastructure could enable the MLPSFSP to access and analyze administrative data, including from the Population Register and Tax Income systems.** The Data Lake infrastructure is fully operational and EU funds will finance a range of improvements—from user training, to hardware and software upgrades—to expand its capabilities in data extraction, enrichment, transformation, storage, analysis, and visualization. The project's completion is expected by July 2026, and the MJPADT plans to develop tools—such as dashboards and automated reports—that could be especially useful for the MLPSFSP's calculations of relevant indicators.

**The MJPADT's Data Lake infrastructure is expected to mediate between data holders, such as the Tax Administration, and data users such as the MLPSFSP.** The MLPSFSP must first receive access approval from data holders, who will then upload the relevant data into Data Lake. Subsequently, the MJPADT will grant the MLPSFSP access to its secure IT system, enabling it to use the data to generate aggregated statistics—such as the AROP and AROPE-Simplified rates.

**For this approach, the MJPADT should meet one additional requirement.** Since the Data Lake project is financed by the EU, it is important for the MJPADT to find other sustainable sources of funding, to ensure it remains operational in the long run. It is anticipated that the State Budget will allocate resources, as Data Lake is part of the State's cloud services. Moreover, plans are in motion to secure financing from the NRRP after the project's completion.

**Requirements for the MLPSFSP concern both data access and technical capacity.** The MLPSFSP should obtain approval from the Tax Administration to access anonymized microdata from the Population Register and Tax Income systems. Moreover, it should ensure the availability of a technical team proficient in statistical and econometric modeling, and of a dedicated monitoring and evaluation team, as described in section 6.2.1.

**This approach capitalizes on the forthcoming secure data processing capabilities of the MJPADT.** By using the Data Lake infrastructure, the MLPSFSP can avoid substantial investments to establish its own secure IT system. Moreover, this approach aligns with the expected requirements around the use of an interoperability platform for data exchange, as outlined in the draft Amendments to the Law on State Data Infrastructure. This not only ensures compliance with upcoming regulations, but streamlines the process of data sharing and collaboration between government entities, enhancing efficiency and safeguarding data integrity.

**Challenges include access to detailed microdata, the recruitment of a capable technical team, and the timely deployment of the MJPADT's Data Lake system.** As discussed in section 6.2.1, it is unclear whether and how the Tax Administration will permit access to its data by other government entities.

---

90 https://esavjetovanja.gov.hr/ECon/MainScreen?entityId=26663

Moreover, as the MLPSFSP would perform all data processing, investment in a strong technical team is crucial. The Data Lake system's enhanced features that are relevant to the MLPSFSP's activities, such as increased data storage capacity and updated software, may not be operational according to the planned schedule, and any delays could affect the MLPSFSP's data processing activities and timelines. Moreover, it is recommended that the MLPSFSP maintain open communications with the MJPADT, to confirm that State Budget financing for Data Lake's long-term operations is secured.

### 6.2.4. Joint implementation by MLPSFSP and CBS

**The CBS is already authorized to use administrative data—including from the Tax Income and, in the near future, the Population Register databases—for statistical purposes.** The Law on Official Statistics and the EU Regulation on European Statistics empower National Statistical Offices to access administrative data sources for the purpose of producing official statistics. The CBS already makes extensive use of administrative data in its statistical production, and continuously integrates new administrative sources. For example, the CBS uses data from the Tax Administration's JOPPD form to create wage and employment statistics and annual Financial Statements for producing Structural Business Statistics, and has recently incorporated Tax Income data into the EU-SILC survey. With the Population Register expected to facilitate the 2031 register-based Population Census, the CBS is likely to gain access to its microdata, which can then be used for statistical activities.

**The role of the CBS may take two different forms.** One entails actively executing key activities, the other consists of facilitating them through the provision of data access, statistical expertise, and additional assistance. The subsections below explain the two potential arrangements in detail.

#### *Option A: The CBS calculates the AROP rate, AROPE-Simplified rate, and income-based social exclusion indicators*

**As part of this approach, the CBS would compute the AROP and AROPE-Simplified rates, as well as income-based indicators of social exclusion, using Population Register and Tax Income data.** For its part, the MLPSFSP would estimate the remaining social exclusion indicators or indexes, and manage the analysis and dissemination phases.

**For the CBS to engage in the production of such indicators, they must be recognized in the statistical program.** Specifically, the calculation of register-based poverty and social exclusion measurements would have to be included in the CBS's Annual Plan of Statistical Activities, whose implementation is contingent upon the availability of resources. The CBS has the technical capacity to produce such estimates. However, the measurement of poverty and social exclusion at the NUTS 3 and LAU levels is considered experimental statistics, as they are not part of the European Statistical Program.

**The CBS is fully equipped to meet all requirements concerning data access, technical expertise, and security of IT infrastructure.** This positions the CBS as a strong provider of data processing services to the MLPSFSP. Per Article 12 of the Law on Official Statistics, the CBS is authorized to offer its services to other entities, thereby creating a source of income. The financial cost of calculating the AROP and AROPE-Simplified rates and income-based social exclusion indicators is expected to be covered by the MLPSFSP, which would have to allocate the necessary funds out of its annual budget.

**On the MLPSFSP's side, the technical requirements for data collection and analysis depend on the indicators or indexes that will be selected.** Their calculation may require the use of tools of varying

complexity, from Microsoft Excel to more sophisticated GIS and statistical software. For reporting purposes, MLPSFSP staff should be skilled in data visualization, through Microsoft Excel or other tools such as Shiny-R.

**The key advantage of this approach lies in the high quality and reliability of the CBS's computations.** The CBS already has access to vital administrative microdata, and a competent technical team. Although initially designated as experimental statistics, the register-based poverty and social exclusion measures have the potential to mature over time, and could ultimately become official statistics—as in the Spanish case discussed in section 4.3.3. Moreover, the CBS's IT infrastructure is highly secure, avoiding the need for dedicated investments by the MLPSFSP.

**The primary obstacle to this option is the scarcity of human resources at the CBS, even assuming that the MLPSFSP would fund all direct expenses.** The CBS is already stretched thin, with a limited budget dedicated to the production of official statistics (and an emphasis on EU-mandated statistics). This may create issues concerning the number of staff members available, their level of expertise, and their capacity to manage a heavier workload without compromising the quality and timeliness of ongoing projects. Before taking on additional responsibilities, the CBS is advised to carefully assess its capabilities and limitations. Should the CBS choose to participate in this effort, a well-thought-out resource management strategy will be important. This could entail streamlining processes, enhancing staff skills through training, or potentially expanding the workforce. The challenge for the MLPSFSP would be to carve out sufficient financial resources every year to support the CBS's work.

### *Option B: The CBS acts as a facilitator*

**This approach is similar to the MLPSFSP-led implementation model outlined in section 6.2.1, but with assistance from the CBS.** Specifically, the CBS would provide the Ministry with access to the necessary data, and to secure IT infrastructure. This option is based on the assumption that the Population Register will be adopted as an official source of statistics. It is expected that the CBS will authorize its users access to the Population Register, and thus enable the merging of its data with that from other administrative sources. The terms of access would mirror those currently in place for research purposes, which involve access to a safe room and strict compliance with a predefined procedure.[91]

**The CBS is equipped to manage the initial stages of data preparation, which involve merging the Population Register and Tax Income microdata.** This foundational task lays the groundwork for the complex statistical activities required to estimate the AROP and AROPE-Simplified rates, and income-based social exclusion indicators. The MLPSFSP is expected to finance the CBS's data preparation services.

**Under the CBS's data access policy, only a select group of pre-approved researchers can handle confidential statistical data.** Therefore, it is important for the MLPSFSP to collaborate with one or several such researchers, who can carry out the relevant statistical activities on its behalf—including the calculation of the AROP and AROPE-Simplified rates, and income-based social exclusion indicators. The MLPSFSP would remain responsible for producing less-complex indicators or indexes, and for analysis and reporting. The relevant technical requirements are the same as those described for Option A, ranging from proficiency in the use of Microsoft Excel and/or GIS and statistical software for data collection and analysis, to data visualization skills.

---

91 https://dzs.gov.hr/podaci-za-znanstvene-svrhe/348

**With this approach, the MLPSFSP would not need to directly handle sensitive microdata or invest in secure data infrastructure.** Instead, the CBS would prepare and merge Population Register and Tax Income data within its own secure IT systems. This approach leverages the CBS's authorization to use administrative data sources and its established security measures. Therefore, it allows the MLPSFSP to avoid the complexities of data access requests and costs associated with setting up and maintaining a secure data infrastructure.

**On the other hand, this approach entails major reliance on external researchers for important statistical activities.** The researchers pre-approved by the CBS would be tasked with complex data analysis and econometric modeling, potentially with limited oversight of their methods and results from the MLPSFSP. This dependency may raise concerns about the quality and integrity of the work conducted. This approach also requires establishing a research contract between the MLPSFSP and the CBS.

# 7. MONITORING SYSTEM FOR POVERTY AND SOCIAL EXCLUSION

**The MLPSFP is advised to establish by 2028 a comprehensive system to monitor poverty and social exclusion.** Table 20 outlines a recommended timeline, from the completion of the Population Register to the establishment of a full-fledged monitoring system by the MLPSFSP.

**Table 20.** Timeline for the development of a monitoring system to track poverty and social exclusion

| Milestone | Responsible stakeholder | Expected timeline |
|---|---|---|
| Completion of the Population Register | Tax Administration | Mid-2026 |
| Estimates of AROP and AROPE-Simplified rates based on Population Register and Tax Income data | MLPSFSP/Tax Administration/ MJPADT/CBS (*) | End-2026 |
| Estimates of social exclusion indicators or indexes | MLPFSP in coordination with data holders (**) | 2025-2026 |
| Complete monitoring system to track poverty and social exclusion | MLPSFSP | 2027-2028 |

*Source: World Bank elaboration.*

*Note: (*) Depending on the institutional arrangements discussed in Chapter 6.*

*(**) See Table 15 for relevant data holders.*

**The proposed monitoring system entails collecting, processing, and reporting on poverty and social exclusion indicators at the NUTS 3 and/or LAU levels.** The MLPSFSP will execute all three activities, with the collection phase also involving other stakeholders to calculate indicators and indicator components, i.e., data holders, using the Population Register and other administrative data sources. The degree of involvement of each stakeholder will depend on the selected indicators and the institutional framework to measure poverty and social exclusion, as described in Chapter 6.

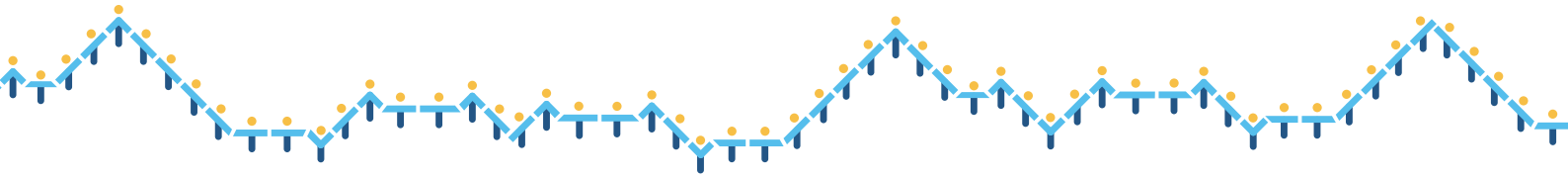
**It is recommended that the MLPSFSP collects, processes, and reports on poverty and social exclusion indicators at the subnational level on an annual basis.** Data collection could take place in January and February of each year for the previous year, while the MLPSFSP could process the indicators by April. On the reporting side, the MLPSFSP is advised to publish the indicators by May. The following sections examine the organization of specific activities within the collection, processing, and reporting phases.

**Adhering to this timeline can help align the monitoring of poverty and social exclusion with the monitoring of national and action plans on social policies[92] within the remit of the MLPSFSP.[93]** In this scenario, up-to-date poverty and social exclusion indicators can readily inform discussions about the impacts, outcomes, and achievements of ongoing social policy measures. Notably, the indicators could highlight changes in the overall social outlook, and help identify ways to better address the needs of vulnerable population groups and territories.

---

92 A proposal for a monitoring system for national and action plans on social policies, including tools for data collection and processing, is delivered under Component 2 of this project, "Development of a methodology for social policies monitoring system with a focus on stakeholder engagements."

93 Ongoing national and action plans with measures to tackle poverty and social exclusion include: 1) the National Plan for the Fight against Poverty and Social Exclusion for 2021-2027, with an Action Plan for 2021-2024; 2) the National Plan for the Equalization of Opportunities for Persons with Disabilities for 2021-2027, with an Action Plan for 2021-2024; 3) the National Plan for the Development of Social Services for 2021-2027, with an Action Plan for 2021-2024; 4) the National Plan for Children's Rights for 2022-2026, with an Action Plan for 2022-2024; and 5) the National Plan for Combating Sexual Violence and Sexual Harassment for the period until 2027, with an Action Plan for the period until 2024.



**In the long run, the MLPSFSP is encouraged to integrate the monitoring of poverty and social exclusion with the monitoring of strategic planning acts on social policies.** For example, poverty and social exclusion indicators could become part of the social policy indicator framework, e.g., as outcome and result indicators for strategic planning acts. This approach can be applied when preparing new strategic planning acts on social policies, and when redesigning ongoing acts. Integrated monitoring can help streamline the collection, processing, and reporting of all indicators relevant to planning, implementing, monitoring, and evaluating policies to address poverty and social exclusion. In turn, the MLPSFSP will be able to directly link poverty and social exclusion indicators to the policy measures enacted and their results.

**The MLPSFSP should aim to put in place an IT system dedicated to monitoring poverty and social exclusion**. An advanced IT solution optimizes data management and mitigates the challenges of collecting data by email and/or official correspondence, which can be burdensome and entails a risk of data loss. The collection and processing form developed as part of this project, presented in spreadsheet format (Excel), could serve as the basis for designing the future IT system—namely, to establish collection protocols and processing rules, as well as presentation formats for reporting purposes. A dedicated IT system should be launched by 2027- 2028, following the operationalization of the data system which enables the estimates of AROP and AROPE-Simplified rates, and other social exclusion indicators.

**Key features of the IT-based monitoring system for poverty and social exclusion should include:**

- Flexibility to incorporate changes to the indicator framework, e.g., to include additional domains and indicators;
- Possibility for multiple users to simultaneously access and enter data in the system;
- Capacity to collect, store, and automatically process large amounts of data, whereby collection refers to data calculated for indicators or indicator components by data holders and processing includes analyzing indicators based on predefined rules. Where applicable, processing incorporates estimating indicators using the data calculated and delivered by data holders for indicator components;
- Ability to produce structured reports and charts;
- Interoperability with other monitoring systems already in place or development (see below for more details).

**The new system should be interoperable, at a minimum, with the MLPSFSP's monitoring system for strategic planning acts on social policies, and with the MRDEUF's strategic planning information system.** Several poverty and social exclusion indicators could already be monitored within the framework of strategic planning acts on social policies. Therefore, interoperability between the relevant systems can enable more efficient monitoring of poverty and social exclusion, while reducing the administrative burden for data holders. Moreover, interoperable systems can help ensure consistency of data over time and across the public administration, while providing additional context (e.g., target values for indicators). While the MLPSFSP plans on using an Excel-based form, similar to the above-mentioned one, to monitor the implementation of national and action plans on social policies, this format is set to be replaced by a more advanced IT solution, linked to the strategic planning information system that the MRDEUF aims to put in place by the end of 2025—to be used for monitoring the preparation of strategic documents, the implementation and evaluation of public policies, impact assessments, and the presentation of progress in implementation. Finally, interoperability should not be limited to the two abovementioned systems: instead, while designing an IT-based monitoring system for poverty and social exclusion, the MLPSFSP should explore ways to streamline the exchange of indicators with a broader range of data holders.

## 7.1. COLLECTION OF INDICATORS

**Data holders should share the poverty and social exclusion indicators and/or indicator components within their remit with the MLPSFSP.** A relevant data holder is identified for every indicator or component of an indicator within the proposed indicator framework (see Table 21). A data holder is an institution tasked with calculating an indicator or component using the Population Register and/or its administrative data sources, and delivering them to the MLPSFSP. A list of institutions identified as data holders is presented in the first sheet of the proposed collection and processing form. For indicators or components based on microdata from the Population Register, the form refers to all three of the institutions considered in Chapter 6, i.e., the Tax Administration, the MJPADT, and the CBS. The sheet also offers an overview of indicators or components within the remit of each institution.

**Table 21.** Overview of data holders for monitoring purposes

| Data holder | Indicator or indicator component |
|---|---|
| **MLPSFSP/Tax Administration/MJPADT/CBS (*)** | AROP rate |
| | AROPE-S rate |
| | Average disposable income |
| | Median disposable income |
| | Average disposable income for the poorest 40% |
| | Average disposable income for the poorest 40%, before social transfers |
| | Population without market income or pensions |
| | Supported population |
| | Population receiving social transfers |
| | Population living off social transfers mostly |
| | Population receiving social transfers as per Law on Social Welfare and Law on Supplement for Inclusion |
| | Employment rate |
| | Population change |
| | Working-age population change |
| | Dependency ratio |
| | Population aged 1-5 |
| | Population aged 15-64 |
| | Population size |
| | Population of children and adolescents |
| | Female population size |
| | Population older than 75 |
| | Households |

| | |
|---|---|
| **Croatian Employment Service** | Number of persons aged 15-64 who participate in the labor market |
| | Number of unemployed persons 15-64 |
| | Number of long-term unemployed persons aged 15-64 |
| **Ministry of Science and Education** | Number of children aged 1-5 enrolled in a kindergarten |
| | Teachers per pupil |
| **National Center for Evaluation of Education** | *Matura* exam performance |
| **Croatian Health Insurance Fund** | Number of medical doctors |
| | Number of dentists |
| **Croatian Health Insurance Fund or Croatian Medical Chamber** | Number of family medicine specialists |
| | Number of pediatricians |
| | Number of school-age and adolescent medicine specialists |
| | Number of gynecologists |
| | Number of psychiatrists |
| | Number of child and adolescent psychiatrists |
| | Number of geriatricians |
| | Addresses of primary health centers |
| | Addresses of hospitals |
| **Croatian Bureau of Statistics** | Area – size of NUTS 3 or LAU |
| | Number of active business entities |
| | Number of active crafts |
| **State Geodetic Administration** | Distance between health center (per address coordinates) and center of the municipality/town, based on GIS data |
| | Distance between hospital (per address coordinates) and center of the municipality/town, based on GIS data |
| **Ministry of the Sea, Transport, and Infrastructure** | Length of paved roads |
| **Croatian Waters** | Length of water pipes |
| **Ministry of Interior** | Number of registered personal cars |
| **Financial Agency** | Number of persons subject to foreclosure |
| **LAUs** | Number of bus lines stopping at a municipality's or town's stations |
| | Number of households connected to sewage |

*Source: World Bank elaboration.*

*Note: (*) depending on the institutional arrangements discussed in Chapter 6.*

**The MLPSFSP should collect indicators and indicator components in a structured and coherent way.** The proposed collection and processing form includes an input form—presented as a separate sheet—developed for every institution identified as a data holder, with the indicators or components within its remit. To facilitate subsequent processing by the MLPSFSP, all input forms follow the same structure and format. From a hyperlink in the list of data holders, in the first sheet of the proposed collection and

processing form, each institution can access the input form for the indicators within its remit. In the future IT-based monitoring system, each data holder may access relevant input forms from institution-specific interfaces, visible upon opening up the system. In addition to the module(s) for the entry of values for indicators or components, each interface should include an overview of all indicators or components under the responsibility of a specific institution, to help data holders prepare for the collection process.

**"Input" refers to indicators or indicator components calculated by data holders.** The current proposal for a poverty and social exclusion indicator framework includes two categories of indicators. One encompasses indicators such as AROP and AROPE-S rates, population receiving social transfers, teachers per pupil, and dependency ratio, among others, to be calculated by data holders using the Population Register and/or their own administrative data sources. The second includes indicators to be estimated by the MLPSFSP, based on components calculated by data holders using the Population Register and/or their own administrative data sources—e.g., unemployment rate, participation rate (see Box 14), density of medical doctors, population density, and density of businesses. In the proposed indicator framework, half of the indicators fall within the second category. Reflecting the categorization of indicators, the first sheet of the collection and processing form, and the input forms for certain institutions, list the components of relevant indicators. Box 14 offers an example of the use of indicator components in the data input and processing form.

**Box 14.** Use of indicator components in data input and processing forms

We consider an example based on the indicator "Participation rate", which refers to the number of people aged 15-64 who participate in the labor market as a percentage of the population aged 15-64. To estimate the participation rate at the NUTS 3 and/or LAU level, the MLPSFSP must first collect data on the number of people aged 15-64 who participate in the labor market and data on the population aged 15-64 in the relevant geographical unit, from—respectively—the Croatian Employment Service (CES), and an institution tasked with estimating indicators based on data from the Population Register (i.e., the Ministry of Finance, the Ministry of Justice, Public Administration, and Digital Transformation, or the Croatian Bureau of Statistics). Therefore, "Participation rate" does not appear on the list of indicators to be calculated and forwarded to the MLPSFSP by data holders, nor in the data input forms for CES, the Ministry of Finance, the Ministry of Justice, Public Administration, and Digital Transformation, or the Croatian Bureau of Statistics. Instead, the indicator list includes the two abovementioned components of the indicator "Participation rate" and the relevant data holders, pointing to the appropriate data input form for each component.

**Upon request of the MLPSFSP, data holders are expected to deliver complete input forms for relevant indicators or indicator components.** The indicators or components should be calculated annually for each NUTS 3 region and/or LAU, and entered into the input form shared with the MLPSFSP through the IT-based monitoring system, respecting the timeline set by the MLPSFSP. Columns in proposed input forms refer to indicators and reporting year, while rows pertain to regions, towns, and municipalities. Access to the system should be administered by the MLPSFSP, and granted to staff from data holders appointed to participate in the monitoring process. The future IT-based monitoring system for poverty and social exclusion indicators should only allow for the submission of forms that contain all requested input.

**Table 22.** Timeline and activities for collecting and processing poverty and social exclusion indicators

| Activity | Output | Timeline* |
|---|---|---|
| **Update input forms within the IT-based monitoring system** | Up-to-date institution-specific input forms available within the IT-based monitoring system | By 15 January |
| – MLPSFSP to update institution-specific input forms used to collect poverty and social exclusion indicators from data holders. Updates should reflect any modifications to the poverty and social exclusion indicator framework | | |
| **Identify persons at data holders responsible for delivering poverty and social exclusion indicators** | List of persons appointed by data holders | By 15 January |
| – MLPSFSP to request data holders to appoint (or confirm the appointment of) person(s) in charge of completing input forms on their behalf | | |
| – Data holders to appoint or confirm persons in charge of delivering poverty and social exclusion indicators | | |
| **Enable access to IT-based monitoring system for data holders** | IT-based monitoring system accessible to appointed persons from data holders | By 15 February |
| – MLPSFSP to grant access to the monitoring system to persons appointed by data holders | | |
| – MLPSFSP to invite appointed persons to complete the institution-specific input forms on behalf of data holders | | |
| **Deliver poverty and social exclusion indicators through the IT-based monitoring system** | Institution-specific input forms submitted by data holders through the IT-based monitoring system | By end of February |
| – Persons appointed by data holders to calculate indicators or components | | |
| – Appointed persons to fill out institution-specific input forms | | |

**All dates fall in the year in which the reporting takes place, e.g., 2025 for reporting on indicators based on data from 2024*

## 7.2. PROCESSING INDICATORS

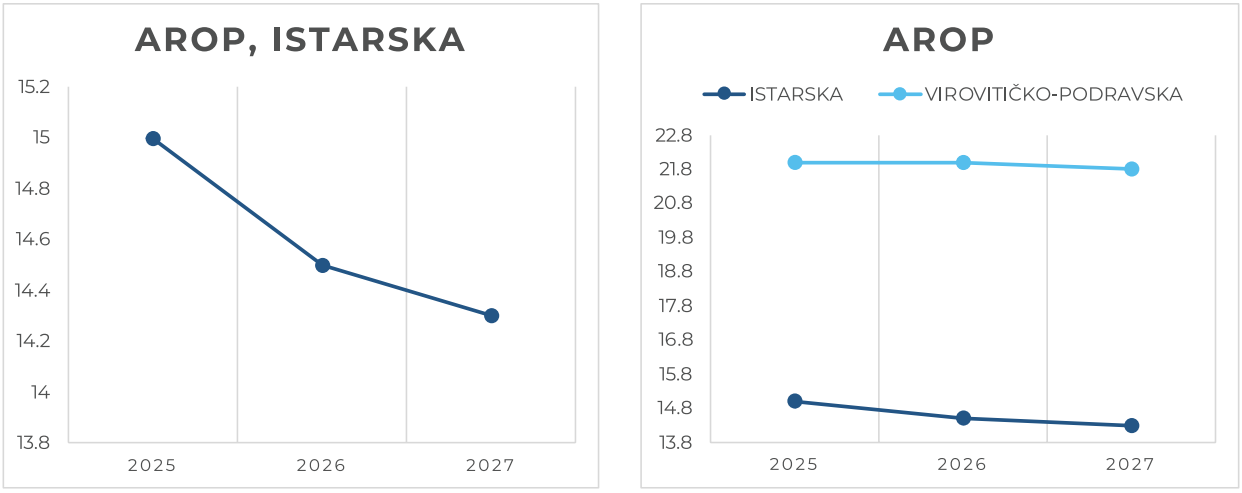
**Indicators are processed to create and present comprehensive information on poverty and social exclusion at the subnational level.** This information is critical to the MLPSFSP's mission of improving social policies and enhancing the transparency, adequacy, and coverage of social benefits. The proposed processing framework includes the following key steps:

- Transferring the input collected into a processing form, developed as part of the IT-based monitoring system for poverty and social exclusion indicators;
- Where applicable, estimating indicators using the data collected for components using the predefined rules;
- Analyzing indicators through graphical representations and select statistics; and
- Creating maps featuring the values of key indicators.

**The proposed processing form enables the automatic transfer of data on indicators and their components from each institution-specific input form.** Indicators calculated by data holders and entered into their respective input forms are transferred instantaneously. For indicators estimated by the MLPSFSP based on data for components delivered by data holders, the processing form is set to perform the estimation automatically drawing on the components' data in the input forms. As a result, the processing form only contains poverty and social exclusion indicators, not their components. Furthermore, the processing form stores indicators for the reporting year as well as for previous years, to enable the analysis of progress over time. The processing form also includes a definition of each indicator, with an explanation of the calculation method for indicators to be estimated by the MLPSFSP based on data on their components.

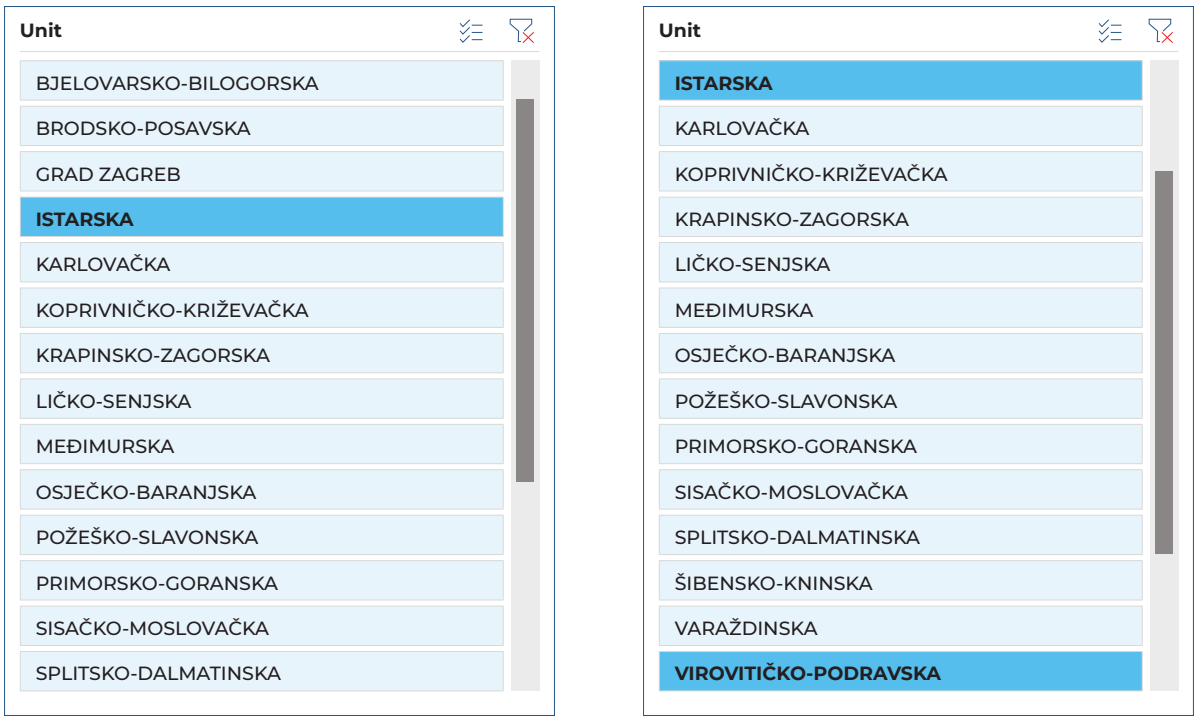
**Graphical representations can be produced in relation to every indicator for one or more NUTS 3 regions and LAUs.** Once the indicators have been processed, the final set of sheets in the proposed collection and processing form generates a variety of graphical representations. One sheet is designed to illustrate progress in the AROP rate, one focuses on the AROPE-S rate, while illustrations for other proposed social indicators are organized by domain, with a separate sheet for each of the nine proposed domains. A graph can be drawn for the trajectory of one indicator over time in one or more NUTS 3 regions or LAUs, as illustrated with simulated data in Figure 10 below.

**Figure 10.** Example of graphical representation for an indicator



A drop-down menu enables the selection of the relevant geographical areas (Figure 11).

**Figure 11.** Selection of administrative unit(s) for the creation of a graph



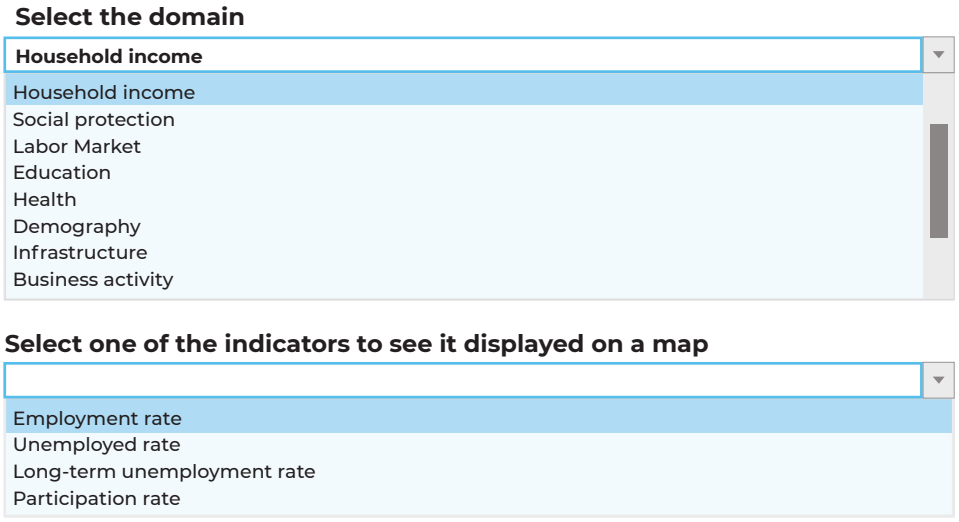

**In a second step of the processing phase, the MLPSFSP could establish, for example, the average, median, highest, and lowest value for each indicator, or the extent of changes in value over time, among other aspects.** To this end, the proposed processing form uses Excel functions to calculate key statistical values. It is recommended that an indicator fact sheet be produced for each poverty and social exclusion indicator, summarizing its selected key values. In the future, such analysis could be conducted through the IT-based monitoring system, setting criteria for automated processing.

**Table 23.** Structure of the proposed processing form

| Sheet | Description of data |
|---|---|
| Overview of trends for: i) all proposed poverty and social exclusion indicators; and ii) all NUTS 3 regions and LAUs | • This sheet contains columns relating to all indicators in the proposed framework. Indicators are grouped by domain. The sheet lists the definition and data source(s) for each indicator, and an explanation of the calculation method for indicators to be estimated based on data on their components. Indicator values for the reporting year and the two previous years are listed in three dedicated columns.<br>• Rows pertain to NUTS 3 regions and LAUs.<br>• The cells are connected with institution-specific input forms, enabling the automatic transfer and/or calculation of indicators. |
| AROP illustration, AROPE-S illustration | • In these two sheets, data on AROP and AROPE-S rates, respectively, is extracted from the processing form for further analysis. Drop-down menus enable the selection of territorial unit(s) for which the illustration of trends is to be produced.<br>• The sheets enable the automatic generation of indicators' fact sheets, including the average, median, highest, and lowest values for each indicator, and the most and least significant changes in values relative to the previous year. |
| Sheets pertaining to social indicators' domains | • These sheets offer the same functionalities described above, but for all indicators within a specific domain. |

**In a third step, indicator values should be plotted on a map, to allow for an easy comparison across geographical areas.** For example, an automated dashboard[94] could build interactive maps based on the selection of a domain and, subsequently, of a specific poverty and social exclusion indicator (see Figure 12). To enable converting data into maps, such a dashboard should be linked to the processing form.

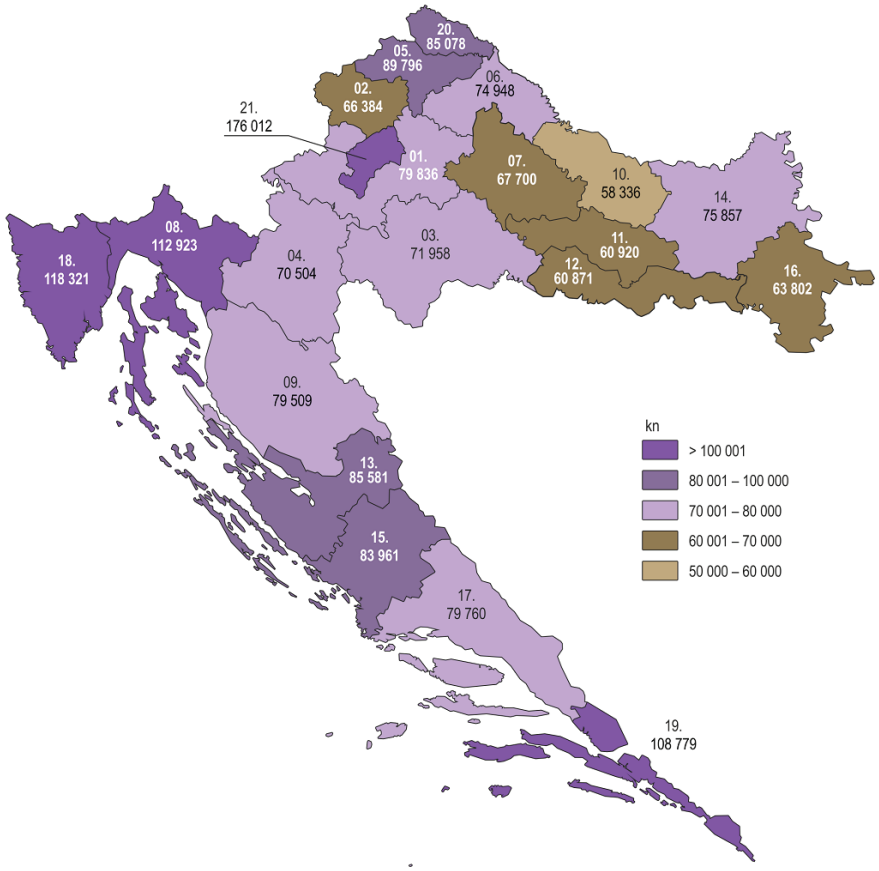
**Figure 12.** Selection of indicator to be visualized on a map




94 See, for example, the application Shiny (https://shiny.posit.co/).

**A map should visualize indicators at the desired administrative level.** Each section of the map, corresponding to a NUTS 3 region or LAU, would be filled with a different color or shade, each reflecting a value range for the selected indicator in the reporting year (as shown in Figure 13). The map should also offer the possibility of displaying indicators for a specific territorial unit by placing the cursor over it. While maps should be created at least for the AROP and AROPE-S rates, the MLPSFSP is invited to use them for other indicators as well, to more easily identify lagging areas in need of attention.

**Figure 13.** Example of visualization of indicator values across subnational units in the reporting year

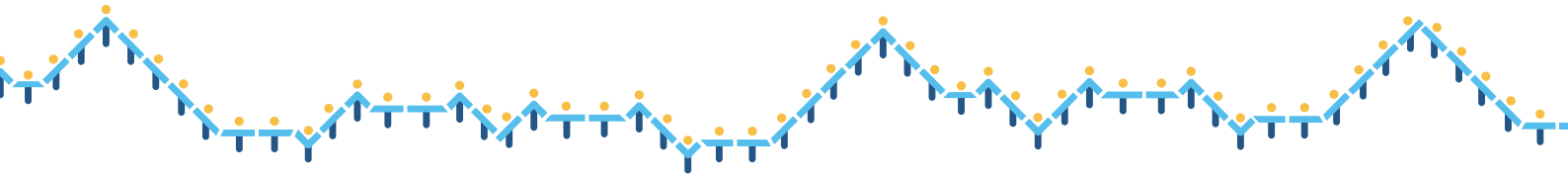


# 7. MONITORING SYSTEM FOR POVERTY AND SOCIAL EXCLUSION

**Table 24.** Step-by-step processing of indicators

| Activity | Output | Timeline* |
|---|---|---|
| **Transfer of inputs into processing form** | Complete the processing form containing all indicators at NUTS 3 or LAU levels for the reporting year | By the end of February |
| – Indicators are automatically transferred into the processing form as data holders enter them into institution-specific input forms.<br>– For indicators based on various components, the processing form automatically calculates relevant values, based on component data delivered by data holders through institution-specific input forms. | | |
| **Analysis of poverty and social exclusion indicators** | Graphical representations of trends for all (or selected) indicators in all (or selected) NUTS 3 regions or LAUs; fact sheet produced for each indicator | By the end of March |
| – MLPSFSP to review trends for all (or some) indicators in all (or some) NUTS 3 regions or LAUs through graphical representations (e.g., trends observed in a specific area, comparison of progress across areas, among others).<br>– MLPSFSP to review values of poverty and social exclusion indicators—e.g., territories with highest and lowest values for each indicator, average and median values of indicators, changes in values, among others. | | |
| **Create maps for poverty and social exclusion indicators** | A map available for all or selected indicators | By the end of April |
| – MLPSFSP to enable the transfer of indicators from the processing form to the selected dashboard tool.<br>– MLPSFSP to define data ranges for drawing maps related to all or selected indicators.<br>– MLPSFSP to create maps. | | |

** All dates fall in the year in which the reporting takes place, e.g., 2025 for reporting on indicators based on data from 2024*



## 7.3. REPORTING ON POVERTY AND SOCIAL EXCLUSION

**Poverty and social exclusion indicators at the subnational level should be published annually, after the relevant data has been updated within the IT-based monitoring system.** Effective communication with the public requires the adoption of user-friendly formats—e.g., a dashboard on the MLPSFSP's website to visualize data on the AROP and AROPE-S rates (as well as other indicators), with a view to enhancing stakeholder engagement, and encouraging public dialogue on addressing the challenges revealed by the data. Reporting on poverty and social exclusion can also play an important role in consultations with the stakeholders during the planning and monitoring of the implementation of social policies. The data should be published along methodological notes (including the definitions of indicators) and a short analysis. The publication should occur by the end of May, to coincide with the MLPSFSP's dissemination of data about progress in the implementation of national and action plans on social policies.

# 8. NEXT STEPS

**The MLPSFSP is advised to adopt a clear roadmap, such as the one suggested in Table 25, to expedite the establishment of a comprehensive system to monitor poverty and social exclusion at the NUTS 3 and LAU levels based on administrative data.** The implementation of the proposed roadmap depends on progress in the development of the Population Register, with the assumption that all concerns mentioned in section 3.7 are resolved satisfactorily. Thus, it is estimated that the monitoring system could be established by 2027-2028 at the earliest.

**Table 25.** Suggested roadmap of next steps

| Step | Responsible stakeholder(s) | Timeline |
|---|---|---|
| Draw up a list of social exclusion indicators that meet the MLPSFSP's needs | MLPSFSP | 2025 |
| Identify important social exclusion topics that administrative data covers insufficiently or not at all | MLPSFSP | 2025 |
| Estimate social exclusion indicators by domain | MLPSFSP and relevant data holders (*) | 2025-2026 |
| Develop and validate statistical model(s) to impute missing income components using Tax Income data | MLPSFSP/ Tax Administration/ CBS (**) | 2025-2026 |
| Design surveys to fill gaps in administrative data | MLPSFSP | 2025-2026 |
| Complete the Population Register | Tax Administration | Mid-2026 |
| Estimate the AROP rate, AROPE-Simplified rate, and income-based social exclusion indicators using Population Register and Tax Income data | MLPSFSP/ Tax Administration/ CBS (**) | End-2026 |
| Develop a geo-referenced database of poverty and social exclusion indicators at the NUTS 3 and LAU levels | MLPSFSP | 2027 |
| Develop an IT system for monitoring that allows for data input from various stake-holders, and produces visual output for reporting and dissemination | MLPSFSP | 2027-2028 |
| Conduct poverty and social exclusion diagnostic at the NUTS 3 and LAU level | MLPSFSP | 2027-2028 |

*Source: World Bank staff elaboration.*

*Note: (*) See Table 15 for relevant data holders.*

*(**) Depending on the institutional arrangements discussed in Chapter 6.*

**It is recommended that the MLPSFSP develop a list of social exclusion indicators that strike a balance between adequacy and practicality.** The MLPSFSP should thoroughly assess which dimensions of social exclusion or inclusion it needs to report on, and with what frequency and level of geographical detail. Moreover, it is crucial to consider the feasibility of the relevant calculations, which can vary greatly depending on technical considerations and the range of stakeholders involved. Ultimately, the chosen indicators should reflect a compromise between the needs of the MLPSFSP, and its capacity to invest financial and human resources to produce them.

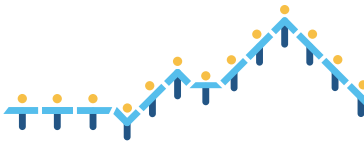
**Since administrative databases may not fully encompass all relevant facets of social exclusion, it is essential for the MLPSFSP to identify topics that require data to be collected via survey, and to prepare for such surveys meticulously.** Conducting surveys on specific topics is generally costlier than extracting data from pre-existing administrative records. Therefore, it is imperative for the MLPSFSP to carefully select survey topics (e.g., homelessness, sexual violence, and issues affecting the Roma community), specify the required level of geographical detail, and determine the frequency of reporting. The MLPSFSP should then meticulously plan for the survey by choosing appropriate data collection methods, developing a sampling strategy (covering sample frame, size, stratification, and re-weighting methodology), and creating the survey's questionnaire as well as guidelines for supervisors and enumerators. The MLPSFSP is also advised to conduct a pilot survey prior to a full-scale rollout. Finally, it is crucial for the MLPSFSP to allocate the budgetary resources necessary to prepare and carry out the surveys.

**It is paramount for the MLPSFSP to engage in discussions with the Tax Administration, the MJPADT, and the CBS to establish an appropriate institutional arrangement for calculating poverty and social exclusion indicators at the NUTS 3 and LAU levels.** The chosen arrangement should be formalized with an agreement that clearly outlines the responsibilities of each party, the respective financial commitments, and the timelines for the execution of various activities. Matters to be covered by the agreement include:

- Data access: the agreement should specify which party will have access to the necessary microdata, the conditions under which access will be granted, and the protocols to ensure the privacy and security of the data.
- IT infrastructure: The agreement should identify the entity responsible for providing secure IT infrastructure to safely exchange, process, and store the data.
- Estimation of indicators: The agreement should clearly allocate the task of estimating crucial indicators, such as the AROP rate, the AROPE-Simplified rate, and other social exclusion indicators or indices.
- Data analysis: The agreement should outline roles and responsibilities for the analysis of the indicators produced, to ensure that they are robust, reliable, and fit to inform policy decisions.
- Visualization and dissemination: The agreement should define who will be responsible for producing visual representations of the findings, to facilitate their interpretation; and for disseminating them to the relevant audience.

**The MLPSFSP is advised to ensure that the agreement clearly addresses financial considerations and timelines related to the production, analysis, and dissemination of indicators.** The agreement should include a detailed breakdown of the costs allocated to each party, and a plan for securing sustainable funding to support these activities over the long term. Finally, it should explicitly define a timeline for the delivery of each component covered by the agreement, accounting for all phases—from data collection to the dissemination of findings.

# 8. NEXT STEPS



**Once the institutional set-up has been agreed upon, the party responsible for estimating the AROP and AROPE-Simplified rates (i.e., the MLPSFSP, Tax Administration, or CBS) is recommended to do so in two phases, reflecting the availability of data from the Population Register.** In the first phase, before the Population Register data is available, the relevant entity should develop an econometric model to estimate missing income components at the individual level. Specifically, the model will use Tax Income data to calculate disposable income in line with Eurostat's definition. To ensure its accuracy, sensitivity tests should be conducted to assess the impact of variations in the model's assumptions or specifications. This phase also entails establishing the AROPE-Simplified status of individuals based on the available Tax Income data. Once the Population Register becomes accessible, the second phase involves merging the estimates of disposable income and AROPE-Simplified status based on Tax Income data with demographic and household information from the Population Register, which in turn makes it possible to calculate the AROP and AROPE-Simplified rates. It is essential that the methodology for both phases is meticulously documented—detailing the procedures adopted, highlighting any methodological deviations from Eurostat's guidelines, and explaining the reasons for them. Comprehensive documentation ensures transparency and allows for future methodological reviews, as new data becomes available or as the understanding of poverty and social exclusion dynamics deepens.

**The MLPSFSP is advised to establish a dedicated IT system for monitoring and analyzing poverty and social exclusion indicators.** Such a system should seamlessly integrate with the MLPSFSP's existing IT infrastructure, which supports various social policy programs, to ensure consistency and efficiency. Furthermore, it should be designed with interoperability in mind, allowing for a smooth exchange of data and for compatibility with the IT system of the MRDEUF. The envisioned IT system would serve as a central repository for a geo-referenced database of poverty and social exclusion indicators at the NUTS 3 and LAU levels, which will enable the MLPSFSP to establish and understand the geographical distribution of poverty and social exclusion. A user-friendly interface would help the MLPSFSP gather data inputs from various stakeholders on an annual basis, ensuring that the information is current and changes can be tracked over time. Moreover, the IT system should offer robust data analysis and visualization capabilities, empowering the MLPSFSP to interpret the data effectively, identify trends, and make informed decisions. Visualization tools would also play a key role in reporting activities, enabling the creation of clear and compelling presentations of the data for a variety of audiences. The MLPSFSP is advised to earmark an annual budget for the development and ongoing maintenance of such a system.

**It is crucial for the MLPSFSP to ensure that the expertise of its staff is commensurate with its wide range of responsibilities.** The Ministry may consider strategically hiring technical staff proficient in econometric modeling and in the use of advanced statistical software—essential skills for estimating poverty and social exclusion indicators. In addition, the MLPSFSP should explore upskilling staff already in charge of monitoring poverty and social policies, by offering structured training programs as well as opportunities for experiential learning—such as peer-to-peer knowledge exchanges and collaborative learning sessions.

# ANNEXES

## A1. AGENDA AND PARTICIPANTS OF STUDY VISITS

### A1.1. Study visit to Ljubljana, Slovenia

The World Bank organized a study visit for 31 Croatian officials to Ljubljana, Slovenia, between January 31st and February 1st, 2024. The Croatian delegation included representatives from the MLPSFSP, the Tax Administration of the MOF, the CBS, the MOJPADT, the MOI, the SGA, and the MRDEUF. Below is the agenda of the study visit.

| Timing | Topic and speaker | Venue |
|---|---|---|
| | **Wednesday 31st January 2024**<br>**Topics: Administrative data system** | |
| 09:30 – 10:15 | Opening remarks:<br>**Margareta Mađerić**, *State Secretary, Ministry of Labour, Pension System, Family, and Social Policy of the Republic of Croatia*<br>**Jana Poljak**, *Head of International Economic Affairs Division, Ministry of Finance of the Republic of Slovenia*<br>**Jehan Arulpragasam**, *World Bank Country Manager for Croatia and Slovenia*<br>**Rima Joujou Deljkić**, *Resident Country Coordinator, DG REFORM*<br>**Ambar Narayan**, *Practice Manager, World Bank*<br>Tour de table introduction | Best Western Premier Hotel Slon |
| 10:15 – 11:15 | **Tomaž Banovec**, *Former Director of the Statistical Office of the Republic of Slovenia*<br>System of registries in Slovenia<br>– Idea,<br>– History<br>– Establishment<br>– Usage | Best Western Premier Hotel Slon |
| 11:15 – 11:30 | Coffee break | Best Western Premier Hotel Slon |
| 11:30 – 12:30 | **Tomaž Banovec**, *Former Director of the Statistical Office of the Republic of Slovenia (continued)*<br>System of registries in Slovenia | Best Western Premier Hotel Slon |
| 12:30 – 14:00 | Lunch | Best Western Premier Hotel Slon |

| | | |
|---|---|---|
| 14:00 – 15:00 | **Ema Pogorelčnik**, *Head of the Building Cadaster Division, Surveying and Mapping Authority of the Republic of Slovenia*<br>Information about apartments in Slovenian building cadastre<br>– Overview of usage, variables, and periodicity of updates<br>– Set-up of the real estate registry<br>– Differences between building cadastre and real estate registry<br>– Census of real estate 2006 | Best Western Premier Hotel Slon |
| **Thursday, 1st February 2024<br>Topics: Usage of administrative data for poverty measures and policies** | | |
| 9:30 – 10:30 | **Ana Božič Verbič**, *Head of Division, Demography and Social Statistics Division, Statistical Office of the Republic of Slovenia*<br>Usage of administrative data for official statistics<br>– Overview of most important sources<br>– Statistical treatment and methodological adaptation related to the use of administrative data sources (including imputation)<br>– Feedback to data holders<br>– Cooperation in the establishment of new administrative sources | Best Western Premier Hotel Slon |
| 10:30 – 11:00 | Coffee break | Best Western Premier Hotel Slon |
| 11:00 – 12:00 | **Stanka Intihar**, *Senior Adviser for Measuring Poverty and Inequality, Demography and Social Statistics Division, Statistical Office of the Republic of Slovenia*<br>Poverty and social exclusion measurements in Slovenia | Best Western Premier Hotel Slon |
| 12:00 – 13:30 | Lunch | Best Western Premier Hotel Slon |
| 13:30 – 14:30 | **Danijel Bratušek**, *Adviser in the User Relations Section, Data Publication and Communication Division, Statistical Office of the Republic of Slovenia*<br>System of accessing the administrative data for research purposes<br>– Overview of available data<br>– Rules for access<br>– Possibilities for data merging<br>– Use case examples | Best Western Premier Hotel Slon |
| 14:30 – 15:00 | Coffee break | Best Western Premier Hotel Slon |

| | | |
|---|---|---|
| 15:00 – 16:30 | Panel discussion – use of administrative data and official statistics for policy purposes<br>**Aša Rogelj**, *Director for Regional Development, Ministry of Cohesion and Regional Development of the Republic of Slovenia (using administrative date and official statistics for regional development)*<br>**Gonzalo Carlos Caprirolo Cattoretti**, *Head of Unit for Analysis and Development, Ministry of Labour, Family, Social Affairs and Equal Opportunities of the Republic of Slovenia (using administrative date and official statistics for social polices and programs)*<br>**Marta Gregorčič**, *Head of the Social Policy Department, Institute of Macroeconomic Analysis and Development of the Republic of Slovenia (using administrative data and official statistics for poverty and social exclusion research)* | Best Western Premier Hotel Slon |
| 16:30 – 17:00 | Wrap-up, conclusions, discussion among participants | Best Western Premier Hotel Slon |

## A1.2. Study visit to Tallinn, Estonia

The World Bank organized a study visit for 11 officials from the MLPSFSP, Tax Administration, and the CBS to Tallinn, Estonia, between June 4rd and 6th, 2024. Below is the agenda.

| Timing | Topic and speaker | Venue |
|---|---|---|
| | **Day 1 – Tuesday 4th June 2024** | |
| 09:30 – 10:00 | **Opening Remark**<br>*Mrs.* **Milena Koren**, *Head of Sector, Ministry of Labor, Pension System, Family and Social Policy*<br>*Mrs.* **Nga Thi Viet Nguyen**, *Senior Economist, World Bank*<br>**An overview of the current situation for digital government development**<br>*Mrs.* **Annela Kiirats**, *Head of Unit, Digital Governance Trainings, eGA* | e-Governance Academy, Rotermanni Street 8, Tallinn |
| 10:00 – 11:00 | **Importance of Digital Government central coordination and policy planning to provide sustainability in secure data processing for decision-making**<br>– E-government policies, strategies<br>– Sustainable development of digital government<br>– Key elements in digital government architecture<br>– Digital registers<br>*Mr.* **Arvo Ott**, *Head of Competence Centre, Digital Architecture, eGA / First government CIO of Estonia (1994-2003)* | e-Governance Academy |
| 11:00 – 11:15 | Coffee break | |
| 11:15 – 12:00 | **End-user focused governance**<br>– Organization of information sharing in secure and legitimate manner.<br>– Function of the data exchange environment for decision-making.<br>*Mrs.* **Annela Kiirats**, *Head of Unit, Digital Governance Trainings, eGA* | e-Governance Academy |
| 12:00 – 13:00 | **Importance of digital identity in all government processes**<br>– Creation of digital identity, based on population register<br>– Digital identity solutions<br>– Verification procedure of a person before actions<br>– Guarantee of a person´s identity and actions online for security<br>*Mr.* **Mark Erlich**, *Senior Expert on eID, eGA* | e-Governance Academy |

| | | |
|---|---|---|
| 13:15 – 14:30 | Lunch | |
| 14:30 – 14:45 | Transfer to Ministry of Interior | |
| 14:45 – 16:45 | **Population and civil register**<br>– Principles of data gathering for the register (who´s data is in the register)<br>– Connections, updates and data sharing<br>– Population register law<br>*Mrs.* **Enel Pungas**, *Director of Population Facts Department*<br>*Mrs.* **Mairis Kungla**, *Deputy Director of Population Facts Department* | Ministry of Interior, Pikk 61, Tallinn |
| | Transfer to the hotel | |
| **Day 2 – Wednesday 5th June 2024** | | |
| 09:30 – 11:00 | **Data Governance and Data Science to support Government Decision-Making**<br>– Estonian Digital Government Agenda 2030 and Government Data Strategy<br>– Stakeholders in data governance<br>– Data requirements for data-driven society<br>– Open data<br>*Mr.* **Ott Velsberg**, *Government Chief Data Officer* | Ministry of Economic Affairs and Communications, Suur-Ameerika 1, Tallinn |
| 11:00 – 11:15 | Transfer to Statistics Estonia | |
| 11:15 – 11:30 | Coffee break | |
| 11:30 – 12:00 | **Introduction to Statistics Estonia. Opportunities in using registers to make useful statistics**<br>*Mr.* **Kaido Paabusk**, *Deputy Director General* | Statistics Estonia, Tatari 51, Tallinn |
| 12:00 – 12:30 | **From 70 million data points to population count of 1,3 million - using registers to produce population statistics in Estonia.**<br>*Mrs.* **Terje Trasberg**, *Team Lead, Population and Education Statistics Team* | Statistics Estonia, Tatari 51, Tallinn |
| 12:30 – 13:00 | **Definition of a household, assembling households based on register data**<br>*Mrs.* **Helle Visk**, *Leading Methodologist, Mathematical Statistics Team* | Statistics Estonia, Tatari 51, Tallinn |

| | | |
|---|---|---|
| 13:15 – 14:00 | Lunch | |
| 14:15 – 14:45 | **Measuring poverty – the process of capturing data from registers**<br>*Mr.* **Taivo Gross**, *Team Lead, Data Acquisition Team* | Statistics Estonia, Tatari 51, T allinn |
| 14:45 – 15:15 | **The role of experimental statistics in the production of official (register-based) statistics**<br>*Mrs.* **Marre Karu**, *Team Lead, Experimental Statistics Team* | Statistics Estonia, Tatari 51, T allinn |
| 15:15 – 15:45 | **Using new data sources for statistics (example of using Electricity bills data for Urban Planning)**<br>*Mr.* **Arko Kesküla**, *Data Scientist, Experimental Statistics Team* | Statistics Estonia, Tatari 51, T allinn |
| 15:45 – 16:00 | **Challenges and future outlooks of register-based statistics**<br>*Mrs.* **Terje Trasberg**, *Team Lead, Population and Education Statistics Team*<br>*Mrs.* **Marre Karu**, *Team Lead, Experimental Statistics Team* | Statistics Estonia, Tatari 51, Tallinn |
| 16:00 – 16:15 | Transfer to the Health Insurance Fund | |
| 16:15 – 16:30 | **Overview of Estonian Health Insurance Database**<br>*Mrs.* **Eda Palm**, *Health Insurance Service Manager* | Health Insurance Fund |
| 16:30 – 16:50 | **Data at the Estonian Health Insurance Fund**<br>*Mrs.* **Ksenia Niglas**, *Data Analyst, Department of Analytics* | Health Insurance Fund |
| 16:50 – 17:15 | **E-Prescription System in Estonia**<br>*Mrs.* **Maris Veermae**, *Health Insurance Service.* | Health Insurance Fund |

| Day 3 – Thursday 6th June 2024 | | |
|---|---|---|
| 09:30 – 11:00 | **Data at Estonian Tax and Customs Board**<br>– Estonian Tax System and tax allocation<br>– Data gathering and usage (salary calculation in the register on taxes)<br>– Connection with different stakeholders for data gathering<br>– Data sharing principles (requirements by the law)<br>*Mr.* **Janek Rozov**, *Deputy Director General, Estonian Tax and Customs Board.* | Estonian Tax and Customs Boards, Lõõtsa 8a, Tallinn |
| 11:00 – 11:30 | **Transfer to Ministry of Social Affairs** | |
| 11:30 – 12:15 | **Estonian Family Policy and Family Benefits system**<br>*Mrs.* **Kadri Raid**, *Head of Family Policy, Department of Children and Families* | Ministry of Social Affairs, Suur-Ameerika 1, Tallinn |
| 12:15 – 13:00 | **Poverty, Subsistence Benefit, and Data**<br>*Mrs.* **Kati Nolvak**, *Head of Economic Security* | Ministry of Social Affairs, Suur-Ameerika 1, Tallinn |
| 13:00 – 14:30 | Lunch | |
| 14:30 – 16:00 | **Roundtable discussion, summary of the week, reflection on lessons learnt**<br>**Wrap-up** | e-Governance Academy |

# A2. EXPECTED VARIABLES IN THE POPULATION REGISTER

| Part of registry | Sect -ion | Variable Description | Source |
|---|---|---|---|
| General | 1 | Personal identifier | OIB registry |
| General | 1 | Surname | OIB registry |
| General | 1 | Name | OIB registry |
| General | 2 | Sex | OIB registry / State registries system |
| General | 2 | Date of birth | OIB registry / State registries system |
| General | 2 | Place of birth – city/municipality | State registries system |
| General | 2 | Place of birth – settlement | State registries system |
| General | 2 | Place of birth – a foreign country | State registries system |
| General | 2 | First residence (city/municipality) – municipality of residence of the mother at the time of birth of a person | State registries system |
| General | 2 | First residence (settlement) – settlement of resi-dence of the mother at the time of birth of a person | State registries system |
| General | 2 | First residence (foreign state) – the country of residence of the mother at the time of birth of a person | State registries system |
| General | 2 | Citizenship | State registries system |
| General | 2 | Citizenship (other) | State registries system |
| General | 3 | Marital status | State registries system |
| General | 3 | Type of registered community (marriage, registered partnership) | State registries system |
| General | 3 | Spouse's personal identifier | State registries system |
| General | 3 | Name and surname of the spouse | State registries system |
| General | 3 | Personal identifier of the mother | State registries system |
| General | 4 | Number of live births | State registries system |
| General | 4 | Mother's name and surname | State registries system |
| General | 4 | Mother's country of birth | State registries system |
| General | 4 | Personal identification of the father | State registries system |
| General | 4 | Father's name and surname | State registries system |
| General | 4 | Father's country of birth | State registries system |
| General | 5 | Adoptive parent's / legal guardian's name and surname | State registries system |
| General | 5 | Adoptive parent's / legal guardian's country of birth | State registries system |
| General | 5 | Adoptive parent's / legal guardian's personal identification number | State registries system |
| General | 6 | City/municipality of temporary / permanent residence | Permanent and temporary residence system |

| | | | |
|---|---|---|---|
| General | 6 | Settlement of temporary / permanent residence | Permanent and temporary residence system |
| General | 6 | Temporary / permanent residence Street | Permanent and temporary residence system |
| General | 6 | Temporary / permanent residence number | Permanent and temporary residence system |
| General | 6 | Supplement to the house number of temporary / permanent residence | Permanent and temporary residence system |
| General | 6 | Reason for registration of residence – Work, education, family reasons, other reasons | Permanent and temporary residence system |
| General | 6 | Reason for registration of residence – Work, education, family reasons, other reasons | Permanent and temporary residence system |
| General | 6 | Date of registration of residence | Permanent and temporary residence system |
| General | 6 | Date of departure from residence (date until when valid) | Permanent and temporary residence system |
| General | 6 | Previous residence (municipality) | Permanent and temporary residence system |
| General | 6 | Previous residence (settlement) | Permanent and temporary residence system |
| General | 6 | Previous residence (foreign country) | Permanent and temporary residence system |
| General | 6 | Country of last residence abroad (if you have ever resided abroad) | Permanent and temporary residence system |
| General | 6 | Date of last immigration from abroad | Permanent and temporary residence system |
| General | 6 | Date of first residence permit for foreigners | Permanent and temporary residence system |
| General | 6 | Expiration date of residence permit for foreigners | Permanent and temporary residence system |
| General | 6 | Date of departure abroad | Permanent and temporary residence system |
| General | 6 | Country of departure | Permanent and temporary residence system |
| General | 6 | Expected time of absence outside the Republic of Croatia | Permanent and temporary residence system |
| General | 6 | The reason for going abroad | Permanent and temporary residence system |
| General | 7 | Disability information | Registry of disabled persons |
| General | | Time of birth | State registries system |
| General | | Date of death | State registries system |
| General | | Time of death | State registries system |
| Special | 1 | The highest-achieving education | Ministry of education - e-system |
| Special | 1 | School attendance and the level it attends | Ministry of education - e-system |
| Special | 1 | Place of schooling – city/municipality | Ministry of education - e-system |
| Special | 1 | Place of education – settlement | Ministry of education - e-system |
| Special | 1 | Place of school – street | Ministry of education - e-system |
| Special | 1 | Place of schooling – house number and house number supplement | Ministry of education - e-system |

| | | | |
|---|---|---|---|
| Special | 1 | Place of school – foreign country | Ministry of education – e-system **** |
| Special | 2 | Basis of insurance | Croatian Health Insurance Institute |
| Special | 3 | Status in activity | Croatian Pension Insurance Institute |
| Special | 3 | Activity NACE code | Croatian Pension Insurance Institute |
| Special | 3 | Occupation | Croatian Pension Insurance Institute ** |
| Special | 3 | Position in employment | Croatian Pension Insurance Institute ** |
| Special | 3 | Place of work – city/municipality | Croatian Pension Insurance Institute |
| Special | 3 | Place of work – settlement | Croatian Pension Insurance Institute |
| Special | 3 | Place of work – street | Croatian Pension Insurance Institute |
| Special | 3 | Place of work – house number and house number supplement | Tax Administration – JOPPD system** |
| Special | 4 | Nationality | State registries system – register of voters |
| Special | 5 | Identification number of the apartment of permanent / temporary residence | State Geodetic Administration (Apartment IDs to be assigned) |
| Special | 5 | information on quality of living | State Geodetic Administration |
| Testimony based | 1 | Data on extramarital union or data on informal life partnership | Testimony based |
| Testimony based | 2 | Mother tongue | Testimony based |
| Testimony based | 2 | Faith | Testimony based |
| Testimony based | 3 | Contact information | Testimony based |
| Additional info | | Source of income | Tax Administration – JOPPD system |
| Additional info | | Type of household | Tax administration / data collection on institutional households |
| Additional info | | Place of school – foreign country | Ministry of education – e-system **** |
| Additional info | | Place of work – foreign country | UNKNOWN, probably testimony based |

# A3. IMPUTATION OF EXPENSES FOR SELF-EMPLOYED INDIVIDUALS AND ACCOMMODATION RENGERS SUBJECT TO LUMP-SUM TAXES.

This appendix provides a detailed technical description of the approach used for the imputation of expenses to the self-employed and accommodation renters with missing expenses in the Tax Administration data.

The approach used is the so-called *multiple imputation* approach, where "multiple" refers to the fact that every missing value of expenses is imputed multiple times. The procedure consists of two parts: estimation and imputation. In the estimation part, expenses are modeled as a function of certain other variables using data on the self-employed who pay regular personal income tax. Then, in the imputation part, the estimated model is used to impute the missing expenses to the self-employed and accommodation renters who pay the tax as a lump sum.

Two procedures are considered here. One is the so-called *regression imputation*, and the other is the so-called *predictive mean matching imputation (PMM)*. The two share the estimation part, differing only in the imputation part.

## A3.1. Notation and preliminaries

There are *complete* and *incomplete* observations in the data. The former are the self-employed who pay regular personal income tax, and whose expenses are available. The latter are the self-employed and accommodation renters with missing expenses. Denote the number of complete and incomplete observations by $n_0$ and $n_1$, respectively. The total number of observations is $n = n_0 + n_1$.

Individual observations are indexed by $i = 1, 2, \dots, n$. Denote by $x = (x_1, x_2, \dots, x_n)'$ the vector of expenses-to-revenue ratio (ERR) for all observations, both complete and incomplete. Some of these, pertaining to the incomplete observations, are missing. $x$ is partitioned into the vector $x_o'$ (of dimension $n_0 \times 1$) pertaining to complete observations and the vector $x_m'$ (of dimension $n_1 \times 1$) pertaining to incomplete observations. The partition is written $x = (x_o', x_m')$. The subscripts $o$ and $m$ refer to "observed" and "missing" values of $x$, respectively.

Let there be $q$ predictors and let $Z_i = (z_{i1}, z_{i2}, \dots, z_{iq})$ denote the values of the predictors of $x_i$ for an observation $i$. The predictors are revenues, dummy variables for residence/location, and dummy variables for NACE sectors of activity. Matrix $Z$ collects $Z_i$ vectors across all observations: $Z_i = (z_{i1}, z_{i2}, \dots, z_{iq})'$. It is partitioned into the matrices $Z_o$ (of dimension $n_0 \times q$) and $Z_m$ (of dimension $n_1 \times q$). The partition is written $Z = (Z_o, Z_m)$.

ERR is assumed to be generated by a normal linear regression model written as

$$x_i | Z_i \sim N(Z_i'\beta, \sigma^2) \quad (1)$$

where $N(\cdot,\cdot)$ stands for normal distribution, $\beta = (\beta_1, \beta_2, \dots, \beta_q)$ is a vector of q unknown regression coefficients corresponding to the $q$ predictors and $\sigma^2$ is an unknown variance.

## A3.2. Regression imputation

Given the above notation and setting, the regression imputation is performed through the following stepwise procedure.

**Step 1.** The linear regression model $x_i|Z_i \sim N(Z_i'\beta, \sigma^2)$ is fit on the data $(x_o, Z_o)$, that is the observations for which expenses are available. In this way, the estimates of the parameters $\beta$ and $\sigma^2$ are obtained. Denote the estimates by $\beta_{est}$ and $\sigma^2_{est}$, respectively.

**Step 2.** New parameters, $\beta_*$ and $\sigma^2_*$, are obtained by simulation under the prior that the probability of $\beta$ and $\sigma^2$ is proportional to $1/\sigma^2$. They are simulated in two stages:

1. $\sigma^2_* \sim \sigma^2_{est} \cdot (n_0 - q) \cdot \chi^{-2}_{n_0-q}$, where the distribution is Scaled Inverse Chi-squared distribution with $n_0 - q$ degrees of freedom;
2. $\beta_*|\sigma^2_* \sim N(\beta_{est}, \sigma^2_* \cdot (Z_o'Z_o)^{-1})$.

**Step 3.** One set of imputed values, $x^1_m$, is obtained by simulating from the normal distribution $N(Z_m\beta_*, \sigma^2_* \cdot I_{n_0 \times n_0})$, where $I_{n_0 \times n_0}$ is an $n_0 \times n_0$ identity matrix (ones on the diagonal, zeros off the diagonal).

**Step 4.** Steps 2 and 3 are repeated $M$ times to obtain $M$ sets of imputed values: $x^1_m, x^2_m, \ldots, x^M_m$.

**Step 5**. Multiply the imputed ERRs $x^1_m, x^2_m, \ldots, x^M_m$ by the revenues to obtain the corresponding expenses $e^1_m, e^2_m, \ldots, e^M_m$.

Steps 1 to 4 can be implemented using the Stata command mi impute regress.

## A3.3. PMM imputation

Steps 1, 2, 4, and 5 are identical to those in regression imputation. The only difference is in Step 3, which is here performed as follows.

**Step 3.** Denote by $x_i^{pred}$ the linear prediction of $x_i$ based on the predictors $Z_i$ and the simulated parameters $\beta_*$: $x_i^{pred} = Z_i'\beta_*$. This prediction is made for all observations (both complete and incomplete). Then, for every incomplete observation $i$, a selected number ($k$) of complete observations (indexed by $j$ ) with the smallest absolute distance $\left|x_i^{pred} - x_j^{pred}\right|$ is found. Call these $k$ observations the nearest neighbors of $i$ and denote their indices by $j = j_1^{min}, j_2^{min}, \ldots, j_k^{min}$. Then one of the nearest neighbors is randomly chosen and its actual value of $x_j$ (not the predicted one, $x_j^{pred}$) is imputed to the incomplete observation $i$ . Repeating this for all incomplete observations gives one set of imputed values, $x^1_m$.

Steps 1 to 4 can be implemented using the Stata command mi impute pmm. The procedure can be done for different numbers of nearest neighbors, say $k$ = 1, 5, 10.

## A3.4. Comparing regression and PMM imputation

Regression imputation is a fully parametric procedure. Both the estimation and imputation parts are parametric. The estimation part is parametric in that it relies on a parametric model of $x$, namely the normal linear regression model. The imputation part is parametric in that the missing values of $x$ are replaced by the values predicted based on the normal linear regression used in the estimation part.

PMM imputation is only partly parametric (semi-parametric). Only the estimation part is parametric, based on the same parametric model as in regression imputation. The imputation part is non-parametric, however. Once the set of nearest neighbors is determined based on parametrically predicted values of $x$, the missing values are replaced by values observed among the nearest neighbors, rather than parametrically predicted.

The parametric nature of the imputation part of regression imputation has consequences for the way missing values are imputed. The normal linear regression model used in the estimation part does not restrict the predicted values to fall within certain sensible bounds of the variable $x$. Concretely, in the present case, where $x$ is ERR, zero is the natural lower bound since expenses cannot be negative. However, there is nothing to prevent the prediction for an observation $i$, $x_i^{pred} = Z_i'\beta_*$, to be negative. Thus, regression imputation may fill missing expenses with implausible values which should be replaced with plausible values before using the data in analyses. A negative predicted value of expenses suggests that the observation has very low expenses according to the model. Since zero is the natural lower bound, replacing negative predictions with zeros seems reasonable.

This, however, cannot happen with PMM imputation since the values to be imputed are not parametrically predicted as in the case of regression imputation. Only the set of nearest neighbors is determined based on parametrically predicted expenses. Once this set is determined for an observation with missing expenses, what is imputed is the observed (rather than parametrically predicted) ERR of a randomly chosen observation from the set of nearest neighbors. Consequently, the imputed ERR cannot be negative, as the imputed value is one observed in the data. Thus, using PMM imputation ensures that any imputed value is within the range observed in the data.

## A3.5. Assessing the performance of different imputation procedures

The performance of regression and PPM imputations should be assessed before deciding which one to use. The comparison should consider PMM for different numbers of nearest neighbors. This can be done only with complete observations only. The validation approach is the so-called K-fold cross-validation.

Let there be $P$ different imputation procedures, indexed by $p = 1, 2, \ldots, P$. Procedure $p = 1$ could be regression imputation, procedure $p = 2$ could be PMM imputation with 5 nearest neighbors, procedure $p = 3$ could be PMM with another number of nearest neighbors, and so on. Comparison of the imputation performances of the $P$ procedures using K-fold cross-validation is performed through the following steps:

**Step 1.** Take all complete observations and divide them randomly into $K$ groups (folds) of about equal size. $K$ should be chosen so that the folds are large enough to be representative of all complete observations; that is, each fold should be a representative sample of the population of complete observations. The most commonly chosen number of folds is $K = 5$ and $K = 10$. Call the folds $\mathrm{F}_1$, $\mathrm{F}_2$, ..., $\mathrm{F}_\mathrm{K}$ and denote the number of observations in them by $\mathrm{N}_1$, $\mathrm{N}_2$, ..., $\mathrm{N}_\mathrm{K}$, respectively.

**Step 2.** Take one fold, $\mathrm{F}_\mathrm{f}$, $\mathrm{f} = 1,2,\ldots,K$ . Make a duplicate of variable $x$ , call it $x(f,p)$, and set its value to missing for every observation in the fold $F_f$. In this way, for the purpose of validation, fold $F_f$ mimics incomplete observations. The remaining folds, considered together, mimic complete observations.

**Step 3.** Perform an imputation procedure p treating fold $F_f$ as incomplete observations, and the remaining folds as complete observations. This yields $M$ sets of imputations: $x_m^1(f,p)$, $x_m^2(f,p)$, ..., $x_m^M(f,p)$.

**Step 4.** For each of $x_m^1(f,p)$, $x_m^2(f,p)$, ..., $x_m^M(f,p)$, compute the mean squared deviation of the imputed values from the observed values (recall that the observed values are available in variable $x$). Compute the root mean squared errors (RMSEs) $R^1(f,p)$, $R^2(f,p)$, ..., $R^M(f,p)$ between $x$ and $x_m^1(f,p)$, $x_m^2(f,p)$, ..., $x_m^M(f,p)$, respectively. The formula for $R^1(f,p)$ is

$$R^1(f,p) = \sqrt{(1/N_f)\cdot \sum_{\text{all } i \text{ from } F_f} \left(x_{m,i}^1(f,p) - x_i\right)^2}, \quad (2)$$

and similarly for $R^2(f,p)$, ..., $R^M(f,p)$. Average these RMSEs to obtain

$$R(f,p) = (1/M) \cdot \left(R^1(f,p) + \cdots + R^M(f,p)\right). \quad (3)$$

**Step 5.** Repeat Steps 2, 3, and 4 for every fold $f = 1,2,\ldots,K$ to obtain the RMSEs $R(1,p)$, $R(2,p)$,..., $R(K,p)$ and average them to obtain:

$$R(p) = (1/K) \cdot \left(R(1,p) + \cdots + R(K,p)\right). \quad (4)$$

**Step 6.** Repeat steps 2, 3, 4, and 5 for each imputation procedure $p = 1,2,\ldots,P$ to obtain the RMSEs $R(1)$, $R(2)$,...,$R(P)$.

**Step 7.** Compare the $P$ RMSEs obtained in Step 6. The procedure with the smallest RMSD is the best performer.

# A4. ESTIMATING AROP AND INCOME-BASED INDICATORS

## A4.1. Estimating AROP rate after multiple imputation of expenses

Once the best-performing imputation procedure is determined as explained in Section 8.1.5, the AROP rate is estimated on the data that includes the expenses imputed with the best-performing procedure. The estimation is done through the following steps:

**Step 1.** Use the best-performing imputation procedure to generate $M$ sets of imputed expenses, $e_m^1, e_m^2, \ldots, e_m^M$, as in Step 4 of regression and PMM imputation.

**Step 2.** Create $M$ new variables $e^1, e^2, \ldots, e^M$, with missing values for all observations. For complete observations (regular-taxed self-employed), each of these variables will hold the observed expenses. For incomplete observations (lump-sum-taxed self-employed and accommodation renters), each variable will hold the corresponding imputed expenses: $e^1$ will hold the imputed values $e_m^1$, $e^2$ will hold the imputed values $e_m^2$, and so on.

**Step 3.** Use the new expenses $e^1, e^2, \ldots, e^M$, along with other variables, to construct $M$ corresponding household disposable incomes, $y_{disp}^1, y_{disp}^2, \ldots, y_{disp}^M$.

**Step 4.** Based on $y_{disp}^1, y_{disp}^2, \ldots, y_{disp}^K$, compute $M$ corresponding AROP thresholds and AROP rates $AROP^1$, $AROP^2$, ..., $AROP^M$. Average the rates to obtain

$$AROP = \frac{1}{M} \cdot (AROP^1 + AROP^2 + \cdots + AROP^M). \quad (5)$$

## A4.2. Estimating social exclusion indicators based on disposable income after multiple imputations of expenses

Some of the social exclusion indicators defined at the municipality/town level (see Section 5.4.1) are based on household disposable income computed from the Tax Administration data. These are the indicators grouped under the heading “household income.”
In their estimation after multiple imputations of the missing expenses, steps 1, 2, and 3 are the same as in the estimation of the AROP rate presented in the preceding section. The only difference is in step 4, where instead of the AROP rate, another indicator is estimated. Let the indicator be denoted by $I$. Then, step 4 reads as follows.

**Step 4.** Based on $y^1_{disp}, y^2_{disp}, ..., y^M_{disp}$, compute $M$ corresponding estimates of $I$ : $I^1, I^2 ..., I^M$. Average them to obtain

$$I = \frac{1}{M} \cdot (I^1 + I^2 + \cdots + I^M). \tag{6}$$

# A5. ECONOMETRIC ESTIMATION OF INCOME UNDERREPORTED TO TAX AUTHORITIES

## A5.1. Basic setting

Consider two types of taxpayers: employed and self-employed. Let all households be classified as either "employed" or “self-employed” households, based on the income source of the working household members. For simplicity, mixed households are disregarded. Employed households are assumed to report their income truthfully (the reported income equals the true income), whereas self-employed households are assumed to report only a fraction of the true income (the reported income is lower than the true income). The objective is to estimate the fraction of true income that is unreported.

The objective is to estimate the factor $k$, defined by $y^{true} = k \cdot y^{reported}$, which is then converted into the unreported share of true income, that is, the proportion of true income hidden from the tax authority:

$$\frac{y^{true} - y^{reported}}{y^{true}} = \frac{y^{hidden}}{y^{true}} = \frac{k-1}{k}. \tag{7}$$

This share, rather than $k$, is typically reported in the studies using this approach.

The identification of $k$ is based on the relationship between consumption expenditures and income – the so-called Engel curve. To afford a given level of expenditure, both employed and self-employed households need certain amounts of income. For observationally identical employed and self-employed households, these amounts of income are the same. Now, if self-employed households report an income lower than employed households, the difference can be considered an estimate of income underreporting on the part of self-employed households.

The estimation is based on data from a household budget survey. The sample is usually restricted to households with two working-age adults as the only members, where at least the household head is working (employed or self-employed). This is to homogenize the sample, as household composition may be a determinant of expenditures. Another usual sample restriction is to consider only households whose main income source is employment or self-employment income.

Households are categorized as either employed or self-employed in the following way. For a household to be considered self-employed, its main income source must be income from self-employment and at least

one household member must be self-employed. For a household to be considered employed, the main income source must be employment income and neither household member may be self-employed.

The estimates of key parameters are obtained from the Engel "curve" represented by the following linear regression equation:

$$\ln E_i = \alpha + \beta \ln y_i + \gamma D_i + \sum_{s=1}^{S} \phi^s X_i^s + \varepsilon_i, \quad (8)$$

where:

- $i$ is the household index;
- $\ln E_i$ is the natural logarithm expenditures on food;
- $\ln y_i^{inst}$is the natural logarithm of permanent income (instrumented as explained below);
- $D_i$ is a dummy variable equal to one if $i$ is a self-employed household, and zero otherwise;
- $X_i^1, \ldots, X_i^S$ are household characteristics (e.g., demographics, location),
- $\varepsilon_i$ is a normally distributed random error term with zero mean;
- $\alpha, \beta, \gamma, \phi^1, \ldots, \phi^S$ are the parameters to be estimated.

In almost all studies using this approach, expenditure on food, rather than total expenditures, is considered. The measurement error is considered smaller for food expenditures than for other consumption items in household budget surveys. Data on food (including non-alcoholic beverages) expenditures are collected from survey respondents through a diary kept by the respondents themselves over a two-week period, while data on other expenditures are based on respondents' memory of the amount spent over a certain preceding period (say, the last three, six, or twelve months). In principle, the longer the recall period, the larger the reporting error. In addition, food constitutes a major share in the budgets of many households.

According to the so-called permanent income hypothesis, consumption is determined by permanent, rather than current income. Current household income consists of a permanent and a transitory component. Determined by human capital and assets, the permanent component is less volatile and thus more anticipated. The transitory component, being determined by temporary factors such as the current activities of household members, is more volatile and thus relatively unpredictable. However, household budget surveys collect data on current income (in the previous calendar year or the last twelve months). Since current income correlates positively with its permanent component, the former measures the latter with an error. Assuming the measurement error is classical (i.e., if the error is additive and independent of the true value), then using current instead of permanent income in the above regression would bias $\beta$ downwards. To avoid the attenuation bias, current income (i.e., permanent income with measurement error) must be instrumented. This is done by specifying the following regression equation

$$\ln y_i = \pi + \psi Z_i + \sum_{s=1}^{S} \zeta^s X_i^s + \xi_i, \quad (9)$$

where:

- $Z_i$ is an instrument for permanent income, usually the education level of the household head;
- $X_i^1, \ldots, X_i^S$ are the same variables as in Equation 8;
- $\xi_i$ is a normally distributed random error term with zero mean;
- $\pi, \psi, \zeta^1, \ldots, \zeta^S$ are the parameters to be estimated.

Based on the parameter estimates, predictions of $\ln y_i$ are made, denoted by $\ln y_i^{pred}$, and used in equation 8 instead of $\ln y_i$. In addition, the predicted regression residuals are obtained, $\xi_i^{pred} = \ln y_i - \ln y_i^{pred}$, and their variance is computed for the self-employed households ($\sigma_{SE}^2$) and employed households ($\sigma_{EMP}^2$).

The estimate of $k$ is computed as the interval estimate $[\underline{k}, \overline{k}]$, where the bounds are[95]

$$\underline{k} = \exp\left(\frac{\gamma}{\beta} - \frac{1}{2}(\sigma_{SE}^2 - \sigma_{EMP}^2)\right), \tag{10}$$

$$\overline{k} = \exp\left(\frac{\gamma}{\beta} + \frac{1}{2}(\sigma_{SE}^2 - \sigma_{EMP}^2)\right). \tag{11}$$

As a special case, the interval estimate collapses to a point estimate if $\sigma_{SE}^2 = \sigma_{EMP}^2$: $\underline{k} = \overline{k} = k = \exp\left(\frac{\gamma}{\beta}\right)$.

## A5.2. Extension 1: heterogeneity in *k*

One extension is to allow $k$ to differ across self-employed households with different characteristics (Cabral, Gemmel, and Alinaghi 2021).

Consider a characteristic with $L$ modalities, indexed by $\ell = 1, 2, \ldots, L$. For example, it could be gender, with two modalities, or region, with as many modalities as there are regions. The characteristic of interest may be one of the characteristics $X_i^1, \ldots, X_i^S$.

The regression equation 1 is modified by putting the interactions of the self-employment dummy with all modalities of the characteristic of interest:

$$\ln E_i = \alpha + \beta \ln y_i + \sum_{\ell=1}^{L} \gamma^\ell D_i C_i^\ell + \sum_{s=1}^{S} \phi^s X_i^s + \varepsilon_i, \tag{12}$$

for households characterized with the modality $\ell$, and zero otherwise, while all else is the same as in equation 1. There are $L$ interval estimates $[\underline{k}^\ell, \overline{k}^\ell]$, where

$$\underline{k}^\ell = \exp\left(\frac{\gamma^\ell}{\beta} - \frac{1}{2}(\sigma_{SE}^2 - \sigma_{EMP}^2)\right), \tag{13}$$

$$\overline{k}^\ell = \exp\left(\frac{\gamma^\ell}{\beta} + \frac{1}{2}(\sigma_{SE}^2 - \sigma_{EMP}^2)\right), \tag{14}$$

or just $\overline{k}^\ell = \overline{k}^\ell = k^\ell = \exp(\gamma^\ell/\beta)$ if $\sigma_{SE}^2 = \sigma_{EMP}^2$.

## A5.3. Extension 2: only public sector employees report truthfully

In this extension, only the incomes of public sector employees are truthfully reported, while those of employees in the private sector and the self-employed may be underreported (Paulus 2015).

The self-employed households are identified in the basic setting: for a self-employed household, the main income source must be income from self-employment and at least one household member must be self-employed. The public-sector employed households are distinguished from the private-sector counterparts. For a household to be considered a public-sector employed household, the main income source must be income from employment, neither household member may be self-employed, and at least one member must be employed in the public sector.

For the implementation of this extension, the regression equation 1 is modified as follows:

$$\ln E_i = \alpha + \beta \ln y_i + \gamma^{SE} D_i^{SE} + \gamma^{EPRI} D_i^{EPRI} + \textstyle\sum_{s=1}^{S} \phi^s X_i^s + \varepsilon_i. \tag{15}$$

---

**95** For derivation, see, for example, Kukk et al. (2020).

Besides the dummy variable indicating self-employed households, here denoted by $D_i^{SE}$ (but otherwise identical to $D_i$ in equation 1), there is a dummy variable indicating private-sector employed households, $D_i^{EPRI}$.

There are now two interval estimates of $k$ : $\left[\underline{k}^{SE}, \bar{k}^{SE}\right]$ and $\left[\underline{k}^{EPRI}, \bar{k}^{EPRI}\right]$. Denoting by $\sigma_{EPRI}^2$ and $\sigma_{EPUB}^2$ the variances of the predicted regression residuals for private-sector and public-sector employed households, respectively, the bounds of the interval estimates are

$$\underline{k}^{SE} = \exp\left(\frac{\gamma^{SE}}{\beta} - \frac{1}{2}(\sigma_{SE}^2 - \sigma_{EPUB}^2)\right), \tag{16}$$

$$\bar{k}^{SE} = \exp\left(\frac{\gamma^{SE}}{\beta} + \frac{1}{2}(\sigma_{SE}^2 - \sigma_{EPUB}^2)\right), \tag{17}$$

$$\underline{k}^{EPRI} = \exp\left(\frac{\gamma^{EPRI}}{\beta} - \frac{1}{2}(\sigma_{EPRI}^2 - \sigma_{EPUB}^2)\right), \tag{18}$$

$$\bar{k}^{EPRI} = \exp\left(\frac{\gamma^{EPRI}}{\beta} + \frac{1}{2}(\sigma_{EPRI}^2 - \sigma_{EPUB}^2)\right), \tag{19}$$

which collapse to point estimates if the variances of the residuals are equal.

# A6. CONSTRUCTION OF INDEX OF MULTIPLE SOCIAL EXCLUSION

## A6.1. Stepwise procedure

The point of departure is the indicators of social exclusion proposed in section 5.4.2. For the purpose of exposition here, let there be $D$ indicators, indexed by $i$ = 1, 2, …, $D$. Let the 555 municipalities and towns in Croatia be indexed by $m$ = 1, 2, …, 555. The value of indicator $i$ for municipality/town $m$ is denoted $I_m^i$.

**Step 1.** For every indicator $i$, compute the correlation coefficient between indicator $i$ and the AROP rate (estimated before). Divide the indicators into two groups: (A) indicators with a positive correlation with the AROP rate; (B) indicators with a negative correlation with the AROP rate. For indicators in group A (group B), a higher value indicates a higher (lower) level of social exclusion.

**Step 2.** For every indicator in group A, rank the municipalities/towns by the value of the indicator in descending order, so that the municipality/town with the highest (lowest) value gets rank 1 (555). For every indicator in group B, rank the municipalities/towns by the value of the indicator in ascending order, so that the municipality/town with the lowest (highest) value gets rank 1 (555). Denote the ranks for indicator $i$, by $R^i$. Then, the rank of municipality/town $m$ for indicator $i$ is denoted $R_m^i$.

**Step 3.** Transform the ranks $R_m^i$ into percentile ranks $P_m^i = \left(\frac{R_m^i}{555}\right) \cdot 100$. The value $P_m^i$ is the percentage of all municipalities/towns with a rank smaller or equal to $R_m^i$, that is the percentage of all municipalities/towns with a higher level of social exclusion by indicator $i$.

**Step 4.** Select a cut-off percentile rank $P^*$ higher than zero and lower than 100. $P_m^i < P^*$ ($P_m^i \geq P^*$) indicates that municipality/town m is considered to have a high (low) level of social exclusion by indicator $i$. For example, $P^*$ could be the median ($P^*$ = 50). Create a variable $LOW_m^i$ equal to one if $P_m^i < P^*$, and zero if $P_m^i \geq P^*$.

**Step 5.** Assign weights $w^1, w^2, \ldots, w^D$ to the indicators, each larger than zero and smaller than one, which add up to one: $\sum_{i=1}^{D} w^i = 1$. An often-used special case is when all weights are equal: $w^1 = w^2 = \cdots = w^D = \frac{1}{D}$. The weight of an indicator measures the importance given to it, relative to the other indicators.

**Step 6.** For every municipality/town $m$, compute the 'average social exclusion', $ASE_m$:

$$ASE_m = \sum_{i=1}^{D}\left(w^i \cdot LOW_m^i\right) = \frac{1}{D} \cdot \sum_{i=1}^{D} LOW_m^i, \tag{20}$$

where the second inequality is due to $w^1 = w^2 = \cdots = w^D = \frac{1}{D}$. The Index of Multiple Social Exclusion for municipality/town $m$ is equal to $ASE_m$.

## A6.2. Temporal comparisons

As constructed above, IMSE allows only comparisons between municipalities/towns in a given year. This is useful to determine the municipalities/towns lagging behind the others. However, meaningful temporal comparisons cannot be made. The reason is the fact the cut-off $P^*$ selected in Step 4 is a relative value and as such may change over time. Suppose that all municipalities/towns experience improvements on all indicators in such a way that the percentile ranks of all municipalities/towns and all indicators remain the same. IMSE would not change, despite the improvements.

For meaningful temporal comparisons, the cut-off must be anchored to a year $t$. Let the cut-off for that year, $P_t^*$ correspond to the indicator values $I_*^1, I_*^2, \ldots, I_*^D$ in year $t$. For example, if $P_t^* = 50$, then $I_*^1, I_*^2, \ldots, I_*^D$ are the medians of the indicators in year $t$. In year $t$+1, instead of a single cut-off applying to all indicators, there will be a cut-off for each indicator: $P_{t+1}^{*1}, P_{t+1}^{*2}, \ldots, P_{t+1}^{*D}$, corresponding to the values $I_*^1, I_*^2, \ldots, I_*^D$, respectively. For example, consider indicator $i$. Suppose that in the anchor year $t$ the cut-off is $P_t^* = 50$, which means that $I_*^i$ is the median in year $t$. However, in year $t$+1, the value $I_*^i$ need not be the median. Suppose that it is the 40$^{th}$ percentile instead. Then, the cut-off percentile for the indicator $t$ in year $t$ +1 is $P_{t+1}^{*i} = 40$. The cut-offs for all years after year $t$ must be anchored to year $t$.

## A6.3. Alternative weighting schemes

The weights introduced in Step 5 need not be equal across the indicators. Since they reflect the relative importance of indicators, one can choose to give more importance to one indicator, and less to another. This is done by increasing the weight of the former and decreasing the weight of the latter in such a way that the weights remain strictly between zero and one, and that their sum remains equal to one. In this way, one can assign weights to the indicators according to policy priorities. For example, one may wish to focus more on social exclusion factors related to health care, in which case one would increase the weights of the health-care-related indicators (e.g., the number of doctors per 1000 inhabitants) and decrease some other weights accordingly.

A related alternative with respect to the weighting scheme is to divide the indicators into several groups, called 'domains' of social exclusion; say, economy, health care, education, infrastructure, etc. Then, two levels of weights are determined: domain weights and indicator weights. At each level, the weights add up to one. Also, the indicator weights within a domain add up to the weight of that domain. The weights can but need not be equal across the domains. Also, the indicator weights may but need not be equal across the indicators within a domain.

To formalize, suppose there are $\mathbb{D}$ domains indexed by $d$ = 1, 2, …, $\mathbb{D}$. Let domain $d$ comprise $n(d)$ indicators indexed by $i$ = 1, 2, …, $n(d)$. The weight of domain $d$ is denoted by $w^d$, and the weight of indicator $i$ within domain $d$ by $w^{d,i}$. Thus, $\sum_{d=1}^{\mathbb{D}} w^d = 1$, and, for each $d$, $\sum_{i=1}^{n(d)} w^{d,i} = w^d$. The $ASE_m$ in Step 6 is then calculated as

$$ASE_m = \sum_{d=1}^{\mathbb{D}} \sum_{i=1}^{n(d)} \left(w^d \cdot w^{d,i} \cdot LOW_m^{d,i}\right), \tag{21}$$

where $LOW_m^{d,i}$ is constructed in the same way as in Step 4, except that now it pertains to indicator $i$ within domain $d$, rather than to indicator $i$.

For illustration, suppose there are two domains, 1 and 2, with weights $w^1$ and $w^2$, respectively. Domain A comprises two indicators (with weights $w^{1,1}$ and $w^{1,2}$), and domain B comprises four indicators (with weights $w^{2,1}$, $w^{2,2}$, $w^{2,3}$, and $w^{2,4}$). Then, $w^1 + w^2 = 1$, with $w^1 = w^{1,1} + w^{1,2}$ and $w^2 = w^{2,1} + w^{2,2} + w^{2,3} + w^{2,4}$. The $ASE_m$ is then calculated as follows:

$$ASE_m = \sum_{i=1}^{2} \left(w^1 \cdot w^{1,i} \cdot LOW_m^{1,i}\right) + \sum_{i=1}^{4} \left(w^2 \cdot w^{2,i} \cdot LOW_m^{2,i}\right). \tag{22}$$

The determination of weights need not be an 'arbitrary' decision as described here. Alternatively, one could rely on a statistical procedure such as the principal component analysis or some other reasonable way to propose a set of weights. Decancq and Lugo (2013) provide an extensive review of the available approaches. To be sure, any approach is 'arbitrary'.

## A6.4. Intensity of deprivation

Variable $LOW_m^i$ created in Step 4 indicates just whether municipality/town $m$ is below the cut-off $P^*$, but not how much below it. Consequently, the $ASE_m$ computed in Step 6 is insensitive to the intensity of social exclusion of municipality/town $m$. For illustration, consider two municipalities/towns where both have an elevated level of social exclusion by any indicator. One is, however, ranked worse than the other by any indicator. Nevertheless, they will have the same value of $ASE_m$ and thus considered to have equal levels of social exclusion. To make the final index sensitive to the intensity of social exclusion, Steps 4 and 6 should be modified as follows.

**Step 4 (modified).** Define a cut-off percentile rank $P^*$ higher than zero and lower than 100. $P_m^i < P^*$ ($P_m^i \geq P^*$) indicates that municipality/town $m$ is considered to have a high (low) level of social exclusion by indicator $i$. Create a variable $GAP_m^i$ equal to the difference $\left(P^* - P_m^i\right)$ for municipalities with $P_m^i < P^*$, and zero for municipalities with $P_m^i \geq P^*$.

**Step 6 (modified).** For every municipality/town $m$, compute the 'average social exclusion gap',

$$ASEG_m = \sum_{i=1}^{D} \left(w^i \cdot GAP_m^i\right), \tag{23}$$

or the 'average social exclusion gap squared',

$$ASEGS_m = \sum_{i=1}^{D} \left(w^i \cdot \left(GAP_m^i\right)^2\right). \tag{24}$$

The Index of Multiple Social Exclusion for municipality/town $m$ is equal to $ASEG_m$ or $ASEGS_m$.

Both $ASEG_m$ and $ASEGS_m$ are sensitive to the intensity of social exclusion, and the latter is more sensitive than the former.

# REFERENCES


Ahmad, Sultan. 2003. "Purchasing Power Parity (PPP) for International Comparison of Poverty: Sources and Methods." World Bank, Washington, DC. https://thedocs.worldbank.org/en/doc/113771487098064473-0050022017/original/AhmadPovertyF.pdf

Albarea, A., M. Bernasconi, A. Marenzi, and D. Rizzi. 2020. "Income Underreporting and Tax Evasion in Italy: Estimates and Distributional Effects." *Review of Income and Wealth* 66 (4): 904-30

Alkire, S., and J. E. Foster. 2011. "Counting and Multidimensional Poverty Measurement." *Journal of Public Economics* 95 (7-8): 476-87.

Amendola, N., G. Mancini, S. Redaelli, and G. Vecchi. 2023. "Price Adjustments and Poverty Measurement." Policy Research Working Paper 10426, World Bank, Washington, DC.

Andrade, C. 2020. "The Limitations of Online Surveys." *Indian Journal of Psychological Medicine* 42: 575-6. https://doi.org/10.1177/0253717620957496.

Angel, S., F. Disslbacher, S. Humer, and M. Schnetzer. 2019. "What Did You Really Earn Last Year? Explaining Measurement Error in Survey Income Data." *Journal of the Royal Statistical Society Series A: Statistics in Society* 182 (4): 1411-37.

Atkinson, A. B., A.-C. Guio, and E. Marlier, ed. 2017. *Monitoring Social Inclusion in Europe, 2017 Edition*. Luxembourg: Publications Office of the European Union.

Australian Social Inclusion Board. 2012. *Social Inclusion in Australia: How Australia is Faring*. Department of the Prime Minister and Cabinet. https://apo.org.au/sites/default/files/resource-files/2012-08/apo-nid30582.pdf

Auten, G., and P. Langetieg. 2023. "The Distribution of Underreported Income: What Can We Learn from the NRP." Manuscript, Internal Revenue Service. https://www.irs.gov/pub/irs-soi/23rpdistributionunderreportednrp.pdf.

Ayala, L., Jurado, A., and Pérez-Mayo, J. 2014. "Drawing the Poverty Line: Do Regional Thresholds and Prices Make a Difference?" *Applied Economic Perspectives and Policy* 36 (2): 309-32.

Bartels, Ch., and Metzing, M. 2019. "An Integrated Approach for a Top-Corrected Income Distribution." *Journal of Economic Inequality* 17: 125-43.

Besim, M., and G. P. Jenkins. 2005. "Tax Compliance: When Do Employees Behave Like the Self-Employed?" *Applied Economics* 37: 1201-08.

Bjorneby, M., A. Altstadsæter, and K. Telle. 2021. "Limits to Third-Party Reporting: Evidence from a Randomized Field Experiment in Norway." *Journal of Public Economics* 203.

Blanchet, T., I. Flores, and M. Morgan. 2022. "The Weight of the Rich: Improving Surveys Using Tax Data." *Journal of Economic Inequality* 20 (1).

Bollinger, C. R., and I. V. Tasseva. 2023. “Income Source Confusion Using the SILC.” *Public Opinion Quarterly* 87 (S1): 542-74.

Bruch, Ch. 2023. “Imputation of Missing Values in Survey Data.” GESIS – Leibniz Institute for the Social Sciences, Mannheim. doi: 10.15465/gesis-sg_en_044

Burricard, C. 2013. “Transition from Survey Data to Registers in the French EU-SILC Survey.” In *The Use of Registers in the Context of EU–SILC: Challenges and Opportunitie*s, edited by M. Jäntti, T. Veli-Matti, and E. Marlier, Chapter 8. Luxembourg: Publications Office of the European Union.

Burkhauser, R. V., N. Hérault, S. P. Jenkins, and R. Wilkins. 2018. “Top Incomes and Inequality in the UK: Reconciling Estimates from Household Survey and Tax Return Data.” *Oxford Economic Papers* 70 (2): 301–26.

Cabrall A. C. G., N. Gemmell, and N. Alinaghi. 2021. “Are Survey-Based Self-Employment Income Underreporting Estimates Biased? New Evidence from Matched Register and Survey Data.” *International Tax and Public Finance* 28: 284–322.

Carlsson, F; and Holmberg, A. 2008. “Availability, Infrastructure, Use and Reuse of Administrative Data in Statistical Production – A Scandinavian Example.” Paper prepared for the “Seventh Management Seminar for the Heads of National Statistical Offices in Asia and the Pacific,” Shanghai, October 13–15.

Carranza, R., M. Morgan, and B. Nolan. 2023. “Top Income Adjustments and Inequality: An Investigation of the EU-SILC.” *Review of Income and Wealth* 69 (3): 725-54.

Chang, T., and N. Vowles. 2013. “Strategies for Improving Data Reliability for Online Surveys: A Case Study.” *International Journal of Electronic Commerce Studies* 4: 121-30. https://doi.org/10.7903/IJECS.1121.

Chen, X., R. Mungai, S. Nakamura, T. Pearson, A. Wambile, and N. Yoshida. 2020. “How Useful is CPI Price Data for Spatial Price Adjustment in Poverty Measurement? A Case from Ghana.” Policy Research Working Paper 9388, World Bank, Washington, DC.

Christian, C. W. 1994. “Voluntary Compliance with the Individual Income Tax: Results from the 1988 TCMP Study.” *IRS Research Bulletin* 1993/1994: 35–42.

Croatia, Ministry of Regional Development and EU Funds, and World Bank. 2016. *Croatia – Small Area Estimation of Consumption-Based Poverty (Poverty Maps)*. https://razvoj.gov.hr/UserDocsImages/Istaknute%20teme/Kartom%20siroma%C5%A1tva/Croatia%20Small-Area%20Estimation%20of%20Consumption-Based%20Poverty%20(Poverty%20Maps).pdf

Croatia, Ministry of Regional Development and EU Funds, and World Bank. 2017. *Croatia – Index of Multiple Deprivation – Conceptual framework for identifying lagging municipalities and towns in Croatia*. https://razvoj.gov.hr/UserDocsImages/O%20ministarstvu/Regionalni%20razvoj/Index%20of%20Multiple%20Deprivation%20-%20Conceptual%20framework_22_10_2018.pdf

Croatian Bureau of Statistics. 2022. *Quality Report for EU-SILC 2021*. https://ec.europa.eu/eurostat/web/income-and-living-conditions/quality

Dawber, J., Wurz, N., and Smith, P. 2019. "The Development of Regional CPIH: Assessment of Small Area Estimation for Regional Expenditure Weights." Office for National Statistics. https://www.ons.gov.uk/economy/inflationandpriceindices/methodologies/thedevelopmentofregionalcpihassessmentofsmallareaestimationforregionalexpenditureweights.

Deaton, A., and Zaidi, S. 2002. "Guidelines for Constructing Consumption Aggregates for Welfare Analysis." LSMS Working Paper No. 135.

Deaton, A., and Aten, B. 2017. Trying to Understand the PPPs in ICP 2011: Why are the Results so Different?. *American Economic Journal: Macroeconomics*, 9 (1), 243-264.

DeBacker, J., B. Heim, A. Tran, and A. Yuskavage. 2020. "Tax Noncompliance and Measures of Income Inequality." *Tax Notes Federal* 166: 1103–18.

Decancq, K., and M. A. Lugo. 2013. "Weights in Multidimensional Indices of Wellbeing: An Overview." *Econometric Reviews* 32: 7–34.

Duffy, B., K. Smith, G. Terhanian, and J. Bremer. 2005. "Comparing Data from Online and Face-to-Dace Surveys." *International Journal of Market Research* 47: 615-39. https://doi.org/10.1177/147078530504700602.

Early, D.W., and E.O. Olsen. 2011. "Geographical Price Variation, Housing Assistance and Poverty." In *The Oxford Handbook of the Economics of Poverty*, edited by Philip M. Jefferson. Oxford: Oxford University Press.

European Commission. 2018. "Territorial Manifestation of Poverty in the 2014-2020 Cohesion Policy." DG REGIO Policy Note.

European Commission. 2022a. *Methodological Guidelines and Description of EU-SILC Target Variables. 2022 Operation (Version 7).*

European Commission. 2022b. *Special Eurobarometer 529 – Fairness, Inequality, and Inter-generational Mobility*. https://europa.eu/eurobarometer/surveys/detail/2652

European Commission. 2023a. *2023 Country Report – Croatia*. https://economy-finance.ec.europa.eu/document/download/ca39aba8-9b5d-4c6c-b3d0-6bf365a6ec47_en?filename=HR_SWD_2023_611_en.pdf

European Commission. 2023b. *EU Comparative Quality Reports*. https://ec.europa.eu/eurostat/web/income-and-living-conditions/quality

Evans, J., and A. Mathur. 2005. "The Value of Online Surveys." *Internet Research* 15: 195-219. https://doi.org/10.1108/10662240510590360.

Evans, J., and A. Mathur. 2018. "The Value of Online Surveys: A Look Back and a Look Ahead." *Internet Research* 28: 854-87. https://doi.org/10.1108/IntR-03-2018-0089.

Flachaire, E., N. Lustig, and A. Vigorito. 2021. "Underreporting of Top Incomes and Inequality: An Assessment of Correction Methods Using Linked Survey and Tax Data." In 36th International Association for Research in Income and Wealth virtual general conference (IARIW), 23-27 August 2021 (pp. 23-27).

Goedemé, T., and L. Zardo Trindade. 2020. MetaSILC 2015: A Database on the Contents and Comparability of the EU-SILC Income Variables [Data File]. doi: https://doi.org/10.7910/DVN/TLSZ4S.

Geodetic Administration of the Republic of Slovenia. 2024. "Materials Prepared for Study Visit to Slovenia – Using Data for Poverty and Social Exclusion Measurement and Related Policies." Ljubljana, January 31–February 1.

Government of the Republic of Croatia. 2023. *National Reform Programme*. https://commission.europa.eu/document/download/fec2933e-154c-45b7-ac6a-51795ca12a78_en?filename=2023_CROATIA_NRP_EN.pdf

Guio, A. C., E. Marlier, and B. Nolan, ed. 2021. *Improving the Understanding of Poverty and Social Exclusion in Europe, 2021 Edition*. Luxembourg: Publications Office of the European Union.

Inglic, Rihard Tomaz. 2007. "Administrative Data and Registers in EU-SILC." Paper presented at the "Seminar on Registers in Statistics – Methodology and Quality," Helsinki, May 21–23.

Jäntti, M., T. Veli-Matti, and E. Marlier, ed. 2013. *The Use of Registers in the Context of EU–SILC: Challenges and Opportunities*. Luxembourg: Publications Office of the European Union.

Johns, A., and J. Slemrod. 2010. "The Distribution of Income Tax Noncompliance." *National Tax Journal* 63: 397–418.

Jolliffe, D. 2006. "Poverty, Prices, and Place: How Sensitive is the Spatial Distribution of Poverty to Cost of Living Adjustments?" *Economic Inquiry* 442: 296–310.

Kleven, H. J., M. B. Knudsen, C. T. Kreiner, S. Pedersen, and E. Saez. 2011. "Unwilling or Unable to Cheat? Evidence from a Tax Audit Experiment in Denmark." *Econometrica* 79.

Kleven, H. J., C. T. Kreiner, and E. Saez. 2016. "Why Can Modern Governments Tax So Much? An Agency Model of Firms as Fiscal Intermediaries." *Economica* 83: 219–46.

Kukk, M., A. Paulus, and K. Staehr. 2020. "Cheating in Europe: Underreporting of Self-Employment Income in Comparative Perspective." *International Tax and Public Finance* 27: 363–90.

Lohmann, H. 2011. "Comparability of EU-SILC Survey and Register Data: The Relationship among Employment, Earnings, and Poverty." *Journal of European Social Policy* 21 (1): 37-54.

Lynn, P., and L. Lyberg. 2022. "Improving the Measurement of Poverty and Social Exclusion in Europe: Reducing Non-Sampling Errors." Eurostat Statistical Working Paper, Eurostat, Luxembourg. https://ec.europa.eu/eurostat/web/products-statistical-working-papers/-/ks-01-21-385

Martinez-Lopez, D. 2013. "The Underreporting of Income by Self-Employed Workers in Spain." *SERIEs* 4 (4): 353–71.

Mavletova, A. 2013. "Data Quality in PC and Mobile Web Surveys." *Social Science Computer Review* 31: 725-43. https://doi.org/10.1177/0894439313485201.

Martin, José María Méndez. 2015. “The Change of Data Sources in the Spanish SILC.” Paper presented at the “EU-SILC Best Practices Workshop,” London, September 16-17.

Moretti, E. 2010. “Poverty, Inequality and Cost of Living Differences.” Discussion Paper Series DP 2010-07, University of Kentucky Center for Poverty Research.

Nordberg Leif, Ulrich Rendtel, and Edin Basic. 2004. “Measurement Error of Survey and Register Income”. In *Harmonization of Panel Surveys and Data Quality*, 65-88. Wiesbaden: Statistisches Bundesamt.

North, Michael S., and Susan T. Fiske. 2013. “Driven to Exclude: How Core Social Motives Explain Social Exclusion.” In *The Oxford Handbook of Social Exclusion*, edited by C. Nathan DeWall. Oxford: Oxford University Press.

Paulus, A. 2015. “Income Underreporting Based on Income-Expenditure Gaps: Survey vs Tax Records.” ISER Working Paper Series No. 2015-15, Institute for Social and Economic Research, University of Essex.

Pissarides, C. A., and G. Weber. 1989. “An Expenditure-Based Estimate of Britain’s Black Economy.” *Journal of Public Economics* 39 (1): 17–32.

Renwick, T. 2009. “Alternative Geographic Adjustments of U.S. Poverty Thresholds: Impact on State Poverty Rates.” Working Paper 2009-11, U.S. Department of Commerce, Census Bureau, Washington, DC.

Republic of Bulgaria and World Bank. 2018. *Poverty Mapping in the Republic of Bulgaria*. Sofia. https://www.nsi.bg/sites/default/files/files/publications/povmap_en.pdf

Roine Jesper, Jonas Vlachos, and Daniel Waldenström. 2009. “The Long-Run Determinants of Inequality: What Can We Learn from Top Income Data?” *Journal of Public Economics* 93 (7-8): 974-88.

Schmutz, F. 2018. “Income Underreporting by the Self-Employed in Switzerland: An International Comparison.” Finanz-Archiv: *Zeitschrift für das Gesamte Finanzwesen* 74 (4): 481-534.

Simler, Kenneth. 2016. “Pinpointing Poverty in Europe: New Evidence for Policy Making.” World Bank, Washington, DC.

Statistical Office of the Republic of Slovenia. 2024. “Materials Prepared for Study Visit to Slovenia – Using Data for Poverty and Social Exclusion Measurement and Related Policies,” Ljubljana, January 31– February 1.

Statistics Austria. 2014. “Documentation on the Register-based Back Calculation of EU-SILC 2008-2011 in Austria.”

Statistics Estonia. 2024. “Materials Prepared for Study Visit to Estonia – Determining Households and Dwellings from Registers using Graphs.” Tallinn, June 5.

The Norwegian Mapping Authority. 2023. “Finding a Utility Unit Number.” https://www.kartverket.no/eiendom/adressering/finn-bruksenhetsnummer-bolignummer

The Norwegian Tax Administration. 2024. “Add or Change a Property Unit Number.” https://www.skatteetaten.no/en/person/national-registry/change/dwelling-number-bolignummer/

Törmälehto, V. M. 2019. “Reconciliation of EU Statistics on Income and Living Conditions (EU-SILC) Data and National Accounts.” Statistical Working Paper, Eurostat, Luxembourg.

Twenge, J. M., R. F. Baumeister, C. N. Dewall, N. J. Ciarocco, and J. M. Bartels. 2007. “Social Exclusion Decreases Prosocial Behavior.” *Journal of Personality and Social Psychology* 92 (1): 56-66. https://doi.org/10.1037/0022-3514.92.1.56.

UNECE (United Nations Economic Commission for Europe). 2022. *Approaches to Measuring Social Exclusion*. Geneva: United Nations.

United Nations, Department of Economic and Social Affairs. 2022. *The Handbook on Management and Organization of National Statistical Systems: 4th Edition of the Handbook of Statistical Organization*.

Van Beuningen, J., and H. Schmeets. 2013. “Social Capital in 2009: An Index for the Netherlands.” *Procedia – Social and Behavioral Sciences* 72: 73-90.

Vellutini, C., G. Casamatta, L. Bosquet, and G. Poniatowski, 2019. “Estimating International Tax Evasion by Individuals.” Taxation Papers, Working Paper No. 76-2019, European Commission. https://taxation-customs.ec.europa.eu/system/files/2019-10/2019-taxation-papers-76.pdf

Wallgren, Anders, and Britt Wallgren. 2014. *Register-Based Statistics: Statistical Methods for Administrative Data. Hoboken*, NJ: John Wiley & Sons.

World Bank. 2024. *Croatia – Systematic Country Diagnostic Update*. Washington, DC: World Bank.

Yonzan, N., B. Milanovic, S. Morelli, and J. Gornick. 2020. “Drawing a Line: Comparing the Estimation of Top Incomes Between Tax Data and Household Survey Data.” Stone Center Working Paper Series 27: 1–50.


Croatia Country Office
Radnička cesta 80/IX, Zagreb
Tel: +385 (01) 2357 222
www.worldbank.org/croatia
WorldBankCroatia