1

rok 2026
ročník 68

Demografie

revue pro výzkum populačního vývoje

Brian Buh
Part-Time, Full-Time, No-Time? Couples' Employment, Education, and First Birth Transitions in the UK

Oana-Maria Cozma – Andrei-Ionuţ Pricop
Household Economic Status and Migration: Evidence from Central and Eastern European Countries

Farhad Yusifov – Elvin Gurbanov
E-Demography as a Platform for Population Data Analysis: A Conceptual Review and the Case of Azerbaijan

Kristýna Dvořáková – Filip Hon
Regionální diferenciace reprodukčního chování v Česku v letech 2015, 2019 a 2023: deskriptivní a shluková analýza

ČESKÝ STATISTICKÝ ÚŘAD

Časopis Demografie je součástí databází **Web of Science** (v Emerging Sources Citation Index) a **Scopus**.

Demografie is a peer-reviewed journal included in the citation database of peer-reviewed literature Scopus and in the Emerging Sources Citation Index (part of the Web of Science Core Collection).

Vážení čtenáři

časopis Demografie vstupuje do svého 68. ročníku. Již třetím rokem v řadě mu byl udělen impakt faktor, který za rok 2024 dosáhl hodnoty 0,6. Zatímco impakt faktor meziročně poklesl, indexy databáze Scopus zůstávají na stabilní úrovni. Velmi pozitivní zprávou je, že roste podíl anglicky psaných textů – v roce 2025 tvořily téměř 60 % obsahu. Tento vývoj zvyšuje relevanci časopisu v mezinárodním kontextu a dává předpoklad pro budoucí růst citovanosti. Vzhledem k vysokém podílu anglicky psaných příspěvku bylo od roku 2025 zrušeno čistě anglické čtvrté číslo, takže všechna vydání jsou nyní jazykové smíšená. V návaznosti na licencování veškerého obsahu od roku 2024 jako „open access" se v roce 2025 stal časopis Demografie součástí databáze Directory of Open Access Journals. Jako články s otevřeným přístupem jsou příspěvky indexovány i v databázích Web of Science a Scopus.

Po delší době bude mít časopis Demografie monotematické číslo. Redakce vypsala výzvu k předkládání článků s velmi aktuálním tématem „Problematika klesající plodnosti ve vyspělých zemích" do čtvrtého čísla tohoto ročníku. Redakční rada zvolila téma zaměřené na analýzu, určení determinantů a důsledků poklesu plodnosti rozvinutých populacích. Věřím, že příspěvky přispějí do této komplexní odborné debaty. Termín zaslání článků byl stanoven na 31. 3. 2026. Bližší informace naleznete na https://csu.gov.cz/aktuality-mono2026.

Jihlava se v termínu od 20. do 22. května stane hostitelem 55. ročníku konference České demografické společnosti. Zájemci o účast mohou své příspěvky prezentovat buď formou přednášky (v českém či slovenském jazyce), nebo prostřednictvím e-posteru (možné i v angličtině). Uzávěrka pro zasílání abstraktů je stanovena na 10. dubna 2026. Pro ty, kteří by chtěli svůj příspěvek publikovat ve formě článku v časopise Demografie, platí termín odevzdání do recenzního řízení 26. června 2026.

Úspěšné vydávání odborného časopisu je výsledkem úzké spolupráce mnoha lidí, kterým bych chtěl na tomto místě poděkovat. Mé poděkování směřuje nejen k členům redakční rady a redakci, kteří zajišťují chod celého procesu. Stejně tak si vážím práce recenzentů, jejichž odborný vhled je zárukou kvality, a v neposlední řadě děkuji samotným autorům, bez jejichž invence a vědeckého přínosu by toto periodikum nemohlo existovat.

Roman Kurkin
předseda redakční rady

Dear readers,

The journal *Demografie* is publishing its 68th volume in 2026. It has received an impact factor for the third year in a row, which reached 0.6 in 2024. While the journal's impact factor decreased year-on-year, its Scopus index scores have remained steady. The share of English-language articles published in the journal continues to grow, which is very good news. In 2025, nearly 60% of the journal's articles were in English. This trend is amplifying the journal's international relevance and provides a basis for further increases in its citation scores. Given the significant amount of English-language content, the journal ceased the practice of devoting the fourth issue of each volume to English-language articles only, and now all its issues are multilingual. After licensing all its content as 'open access' in 2024, *Demografie* was included in the Directory of Open Access Journals database in 2025. As open-access content, these articles are now indexed in the Web of Science and Scopus databases.

For the first time in a while, *Demografie* is planning a thematic issue in 2025. The editorial board has issued a call for papers on a very timely topic: 'The Issue of Declining Fertility in Developed Countries'. The articles will be published in the fourth issue of this volume. The editorial board selected this topic to direct attention to analysing and identifying the determinants and consequences of declining fertility in developed populations. I believe the articles will substantially contribute to the complex professional debate on this topic. The deadline for submitting articles is 31 March 2026. For more information, please visit https://csu.gov.cz/announcements-mono2026.

From 20 to 22 May 2026, the Czech city of Jihlava will host the 55th annual conference of the Czech Demographic Society. Those interested in participating may present their contributions in the form of a conference paper (in Czech or Slovak) or as an e-poster (in Czech, Slovak, or English). The deadline for submitting abstracts is 10 April 2026. Anyone wishing to publish their work as an article in *Demografie* must submit it for peer review by 26 June 2026.

The successful publication of this scholarly journal is the result of close cooperation among numerous people, whom I would now like to thank. I extend my gratitude to the members of the editorial board and to the editorial staff who ensure the smooth running of the entire process. I am also grateful for the work of the reviewers, whose professional insights are a guarantee of the journal's quality. Last but not least, I would like to thank the authors themselves, without whose creativity and scholarly contributions this periodical could not exist.

Roman Kurkin
Chair of the Editorial Board

PART-TIME, FULL-TIME, NO-TIME? COUPLES' EMPLOYMENT, EDUCATION, AND FIRST BIRTH TRANSITIONS IN THE UK

Brian Buh

Abstract

In low-fertility countries, couples' employment patterns may influence the transition to parenthood, potentially moderated by educational attainment. Using data from the UK Household Longitudinal Study (2009–2023), this study investigates how couples' employment status and education relate to the likelihood of having a first child. Results show that among most dual-earner couples, the probability of first birth is similarly high whether both partners work full-time or women work part-time. Conversely, men's part-time work and single-earner couples are associated with a lower likelihood of becoming parents. No strong evidence supports differences in parenthood transition based on couples' educational attainment. While couples with secondary education exhibit a trend where women's part-time work corresponds with higher first birth probabilities, wide confidence intervals indicate substantial variability, limiting firm conclusions about education's moderating role.

Keywords: First Birth, Part-time work, Couples' employment, Understanding Society, Labour force participation
Demografie, 2026, **68(1): 5–24**
DOI: https://doi.org/10.54694/dem.0374

INTRODUCTION

The Total Fertility Rate (TFR) in the United Kingdom (UK) has declined over the last decade from 1.89 in 2009 to 1.48 in 2022 (*Human Fertility Database*, 2026a). This is fuelled by a delay in first birth, from 29.4 years in 2009 to 30.9 years in 2022 (*Human Fertility Database*, 2026b), and an increase in childlessness (*Ermisch*, 2021). A large strand of literature links the continued decline in TFR, the rising age at first birth, and increased childlessness to changes in the relationship between women's employment and parenthood. The longstanding association where women's unemployment presents an opportunity for motherhood has become less pronounced or disappeared (*Doepke et al.*, 2022). Rather, recent empirical work has observed that being employed has become a precondition to motherhood (*Miettinen – Jalovaara*, 2020), and the effect has become stronger throughout Europe (*Alderotti et al.*, 2021).

Part-time employment may offer a solution to balancing work and family obligations, allowing individuals to maintain ties to the labour force and sustain career opportunities, stability, and income. However, the decrease in working hours comes with the trade-off of lower wages, which is, in turn, negatively associated with parenthood (*Jalovaara – Miettinen*, 2013; *van Wijk et al.*, 2021). Nevertheless, fertility decisions are made within the context

1) Austrian Academy of Sciences, Vienna Institute of Demography. Contact: brianbuh2@gmail.com.

of the couple's employment situation. If one partner maintains stable, well-remunerated employment, the wage penalty may diminish, fostering specialisation in family roles, usually by women. Despite this, recent studies indicate within couples, women's absence from employment exhibits the most adverse correlation with the transition to parenthood (*Di Nallo – Lipps*, 2023; *Scherer – Brini*, 2023) and that one partner in part-time work decreases the likelihood of parenthood compared to both working full-time (*Comolli*, 2021).

Like many countries, the UK has seen a notable increase in the number of people continuing their education beyond the secondary level (*Champion*, 2022), driven by the rising educational achievements of women. Consequently, the educational dynamics within couples have undergone notable changes, marked by a growing prevalence of both homogamy and hypogamy, where women are the same or more educated than men (*Esteve et al.*, 2016). However, evidence linking couples' educational attainment to childbirth is mixed and country-context dependent (*Nitsche et al.*, 2018; *Osiewalska*, 2018). While variation in the relationship between employment, education, and first birth has been observed on an individual level (*Kreyenfeld – Andersson*, 2014; *Miettinen – Jalovaara*, 2020), previous studies have not found a relationship between couples' employment status and first birth by educational attainment (*Bueno – García-Román*, 2021; *Comolli*, 2021).

This article aims to explore gendered and socio-economic patterns of paid employment and the transition to parenthood within couples. I ask, *how are couples' employment statuses, when also considering working hours, associated with entry to parenthood?* Additionally, this study examines whether the relationship between couples' employment statuses and parenthood varies by the educational attainment of the couple. I employ the United Kingdom Household Longitudinal Study (UKHLS) years 2009–2023. The UK serves as an apt case study to assess the influence of couples' employment statuses on the transition to parenthood due to its relatively liberalised employment regulations (*Harkness et al.*, 2023). This flexibility enables workers to transition swiftly between jobs, although transferring significant financial risk onto the household. Additionally, stringent work-related criteria for social welfare assistance in the UK, especially for individuals without children, act as an incentive for active labour force participation (*Vegeris et al.*, 2010). Moreover, economic pressure stemming from real wage stagnation adds additional stress to wage earners (*OECD Statistics*, 2024). Recent research on couples' employment and childbirth in the UK has primarily centred around job loss (*Di Nallo – Lipps*, 2023). This article, in contrast, delves into the impact of both men's and women's part-time employment and explores the potential moderating influence of educational alignment within couples.

2. BACKGROUND

2.1 Couples' Employment and Parenthood

Across European contexts, a negative relationship between women's unemployment and childbirth has increased over time (*Alderotti et al.*, 2021). Research rooted in New Home Economics (*Becker*, 1960) primarily examines women's labour force participation through the lens of an economic trade-off between income and time allocated for childrearing. Nevertheless, there are other factors motivating women with partners in paid employment to actively participate in the labour force that might strengthen the positive relationship between women's employment and family formation. This includes having a second income as insurance, work requirements for social welfare benefits, and self-interest: financial independence, career aspiration, professional identity, or personal fulfilment (*Oppenheimer*, 1994).

The prevalence of dual full-time working couples has also increased, surpassing the number of couples where men work full-time and women work part-time, making it the most common arrangement in 2019 (*ONS*, 2022). By 2021, 50.4% of couples were dual full-time employed, compared to 44.1% of dual-employed couples where the women worked part-time. This indicates a national trend towards dual full-time employment, which varies based on family composition.

Paid employment is not a binary state, and individuals may be in paid work but with reduced hours. In the UK, approximately 43.0% of women work part-time (compared to 12.0% of men) (*Warren*, 2022). Childless individuals may choose part-time work for reasons such as work-life balance, personal interests, career exploration, health, or time for education and hobbies (*Warren*, 2004). Others may be pushed into part-time roles due to limited job availability, economic downturns, or occupational segregation (*Cam*, 2012). Part-time employment can support the transition to parenthood by offering flexibility and lower opportunity costs, but it may also hinder it through reduced income and increased uncertainty. Empirical studies have generally found a negative association between part-time work and first births in France, the Netherlands, and the UK (*Dupray – Pailhé*, 2018; *Begall*, 2013; *Schmitt*, 2012), mixed results in Australia (*Laß*, 2020), and no clear pattern in Germany (*Laß*, 2020; *Schmitt*, 2012). These inconsistencies may reflect the individual-level focus of previous research, overlooking the influence of partners' earnings on part-time work choices.

It may be more insightful to examine employment within the couple context. Partners often adopt compensatory patterns, where one partner's reduced or lost employment is balanced by the other's stable, full-time work (*Blossfeld – Drobnic*, 2001). Dyadic research finds that dual-earner couples have the highest likelihood of a first birth (*Jalovaara – Miettinen*, 2013; *Vignoli et al.*, 2012). However, when accounting for working hours, couples where one partner works part-time are less likely to become parents than dual full-time couples in Spain and the US (*Bueno – García-Román*, 2021; *Comolli*, 2021). It remains unclear, though, whether the gender of the part-time worker matters.

Informed by the theoretical framework and prior empirical research, I hypothesise that couples in which both partners work full-time will be more likely to transition to parenthood than those in other employment arrangements (H1a). Given the gendered nature of part-time work in the UK, I expect that within dual-earner couples, men's part-time employment will be more negatively associated with the likelihood of having a first child than women's part-time employment (H1b). Finally, drawing on recent findings highlighting the negative association between women's unemployment and parenthood, I hypothesise that being out of the labour force will be equally detrimental for both men and women when transitioning to parenthood (H1c).

2.2 Couples' Education

In England and Wales, the proportion of individuals over age 16 with a tertiary degree rose by 6.6% between the 2011 and 2021 censuses, driven largely by women's educational gains (*Office for National Statistics*, 2023). This shift has altered the educational alignment of couples: hypergamy (men more educated than women) is declining, while hypogamy (women more educated than men) has become increasingly common (*Clark – Cummins*, 2023). Still, educational homogamy remains the most prevalent pattern among co-residential couples (*Grow – Van Bavel*, 2015). Research suggests that less educated couples tend to transition to parenthood earlier (*Mikolai et al.*, 2018), while highly educated homogamous couples often delay childbearing (*Nitsche et al.*, 2018). Hypergamous couples also tend to have earlier births than hypogamous ones (*Osiewalska*, 2018), indicating that couples with less educated women generally become parents sooner. However, this pattern is not consistently observed in studies focusing on couples' employment (*Bueno – García-Román*, 2021; *Comolli*, 2021). Based on this, I hypothesise that couples in which both partners hold a secondary degree or lower will show a more positive association with the transition to parenthood compared to other educational groups (H2).

2.3 Educational Heterogeneity in the Relationship Between Couples' Employment and Parenthood

Highly educated couples benefit from greater economic resources, job security, and the ability to outsource domestic tasks, all of which can support the transition to parenthood. However, their pathway to parenthood differs qualitatively. A later school-leaving age delays entry into parenthood (*Ní Bhrolcháin – Beaujouan*, 2012), and the time required to establish a stable career may further postpone childbearing. Additionally, they may face higher career penalties for taking parental leave (*Budig et al.*, 2012) and are often more career-oriented (*Gerber et al.*, 2009), making them less willing to interrupt employment for parenthood (*Savelieva et al.*, 2023).

Among the childless in the UK, part-time work is more prevalent among low-educated women than their higher-educated counterparts (*Manning*

– *Petrongolo*, 2008), while it remains uncommon among men (*Warren*, 2022). Part-time roles are often involuntary (*Pech et al.*, 2021) and linked to limited career prospects and lower job quality (*Warren – Lyonette*, 2018). The associated income reduction also varies by educational background.

Unlike individual-level studies that find a negative association between women's unemployment and first birth (*Alderotti et al.*, 2021), dyadic research has not identified significant variation in the relationship between couples' employment and parenthood by educational attainment (*Bueno – García-Román*, 2021; *Comolli*, 2021). However, drawing on individual-level evidence, I hypothesise that among couples where both partners hold tertiary degrees, the negative association between single-earner arrangements with women not employed and the likelihood of a first birth will be stronger than among less educated couples (H3a). Although there is a lack of empirical research on how part-time employment interacts with education to influence parenthood, I theorise that highly educated couples will experience a more negative association between women's part-time employment and first birth, compared to dual full-time couples, than those with lower educational attainment (H3b).

3. METHODOLGY

3.1 Data and Sample

Starting in 2009, the UKHLS (colloquially known as Understanding Society) is a yearly panel survey focusing on households and family issues (*University of Essex*, 2023). This includes three sample groups (Great Britain, Northern Ireland, and an ethnic minority boost), collected using a clustered and stratified probability approach (*Lynn*, 2009). The second wave adds households that continued from the previous British Household Panel Survey, and the sixth wave includes an additional immigrant and ethnic minority boost. Employment, partnership, and education histories can be quickly created using retrospective and prospective data (*Wright*, 2020). All household members above the age of 16 complete the individual questionnaire, and individuals enumerated in wave 1 are followed if they leave the household. The individual response rate in the first wave was 81.8%, with attrition of about 15.0% after the first wave and an average of less than 2.0% for subsequent waves. Understanding Society is generally representative of British society (*Benzeval et al.*, 2020), and post-stratification weights are provided for dealing with survey design, non-response, and attrition (*Platt et al.*, 2020).

At the time of enumeration of households into the panel, each household member aged 16 and above completes a retrospective life history, detailing significant events such as completing school, cohabitation, or changes in main activity status. These life histories are subsequently updated in each follow-up interview, enabling researchers to reconstruct monthly employment and partnership histories for each respondent. However, in cases where household members do not participate in the initial interview at the time of the household's enumeration, their retrospective life histories are unavailable. This makes using retrospective life histories particularly difficult for studying couples, as many partners enter the panel after enumeration. Nonetheless, this study leverages the life histories as they are constructed in each interview to create monthly employment and partnership observations, beginning from the inception of the UKHLS in January 2009 until January 2023.

The sample for this paper comprises all enumerated coresidential childless couples. All women are childless. Some men may have children from previous relationships, which can be difficult to track due to the underreporting of male fertility (*Joyner et al.*, 2011). Couples enter the sample at the month of enumeration into the panel, the formation of their coresidential partnership, or the woman's sixteenth birthday. Couples stay in the sample until twelve months before the first live birth, they separate from their partner, the month of their forty-fifth birthday, or the month of the last interview. The analytical sample does not include same-sex couples, as the aim is to observe gendered aspects of couples' employment.

Information about employment, partnership, and other sociodemographic characteristics is collected in the individual questionnaire. The household grid can be used to match coresidential partners' individual questionnaire responses and life histories. I started by selecting women with valid information about activity status and education because women have more reliable fertility and parity data. However, missing partner data may occur for various reasons, such as the partner not being coresidential, not being present on the day of the interview, declining to participate, or the respondent opting not to disclose information about their partner. To address this, I exclude monthly observations with no provided partner identification number (7,391 monthly observations), cases where the woman's activity status is missing (575 monthly observations), the male partner's activity status is missing (69,274 monthly observations), or either partner's educational information is missing (1,940 monthly observations)[2]. The final sample comprises 87,481 monthly observations, representing 2,701 unique couples. On average, couples are observed for a mean duration of 65 months.

3.2 Measures

Dependent Variable. This study focuses on the transition to parenthood, measuring couples' characteristics 12 months prior to the birth to minimize reverse causality. The dependent variable is binary, coded as 1 if a live first birth occurs within the following 12 months, and 0 otherwise. The dataset includes 1,044 first birth events, representing 38.7% of the couples. The average age of women at first birth in the sample is 30.7 years, slightly higher than the UK average of 30.4 years during the same period (Human Fertility Database, 2023), likely due to the exclusion of births by single mothers.

Couples Employment. I created a categorical variable for couples' employment status with six groups: (1) *Both full-time*; (2) *Woman part-time, Man full-time*; (3) *Woman employed, Man part-time*; (4) *Woman employed, Man not employed*; (5) *Woman not employed, Man employed*; and (6) *Both not employed*. Employment status is based on self-reported main activity and usual weekly working hours (including overtime). Individuals are considered employed if their main status is "Self-employed" or "Paid employment." Full-time work is defined as 30 or more hours per week. Couples fall into the *Both full-time* group if both work full-time. The *Woman part-time, Man full-time* group includes women working part-time and men full-time. Since men's part-time work is less common (4.6%) and dual part-time couples are rare (1.2%), the *Woman employed, Man part-time* group includes all men working part-time paired with women employed either full- or part-time. The remaining categories cover couples where one or both partners are not employed, regardless of hours worked.

Sensitivity analyses (Table A1), based on logistic regression models, indicate that couples in which women engage in carework or homemaking have a higher likelihood of becoming parents than dual full-time employed couples. However, this group represents only a small share of the sample (2.7% of observations). The main analysis therefore excludes inactive employment statuses, as missing information on men's employment disproportionately reduces these cases. Moreover, including interaction terms for these small groups results in insufficient observations within some educational categories, limiting the robustness and reliability of the estimates. The sensitivity analyses are used to assess whether excluding inactive employment statuses affects the substantive conclusions; results indicate that the main patterns observed in the primary models remain unchanged.

Couples' Education. Educational attainment is recorded at the first interview and retrospectively updated in subsequent waves. Given differences in education systems across England, Northern Ireland, Scotland, and Wales (*Schneider,* 2008), and the separate reporting of education attained outside the UK from wave 6 onwards (which I imputed retrospectively), I harmonised these data using the International Standard Classification of Education (ISCED-97). Education was categorised into tertiary (ISCED-97 levels 5–6) and secondary (levels 1–4). Couples were then grouped by highest attainment into: (1) *Both*

2) A significant number of observations were dropped, particularly those with missing male partners' characteristics, potentially biasing the sample. Generally, the women in the removed observations are more likely to be low educated (22.9% compared to 15.2%), more likely to be unemployed (5.1% compared to 3.6%), inactive (12.9% compared to 5.3%), and between the ages of 16-24 (17.0% compared to 13.9%). This indicates that any evaluation of results from couples with non-employed women and low-educated couples must consider the potential bias when interpreting the findings.

tertiary, (2) *Heterogamous*, and (3) *Both secondary*. Sensitivity analysis (Table A1) shows that hypogamous couples (women more educated) are more likely to transition to parenthood than hypergamous couples (men more educated). However, to ensure sufficient first birth events for interaction analyses, heterogamous couples were combined into a single category. Additionally, a binary variable for enrollment in education was included to account for couples where one or both partners are currently studying.

Other controls. All models control for woman's age, partnership status, and couples' country of birth, as these factors are strongly associated with the timing of first births and may also correlate with couples' employment arrangements. Woman's age captures biological constraints and well-established life-course patterns in fertility, with first births concentrated in the late twenties and early thirties. Partnership status (married versus cohabiting) reflects differences in relationship stability, legal protection, and social norms surrounding childbearing in the UK, where marriage remains strongly associated with entry into parenthood. Couples' country of birth (both UK-born, mixed, or both foreign-born) is included to account for heterogeneity in fertility norms, migration-related timing effects, and labour market integration that may influence both employment patterns and fertility behavior. Including these variables helps isolate the association between couples' employment, education, and the transition to first birth.

Table 1 summarizes the sample's descriptive statistics. Most couples have both partners working full-time, though part-time or non-employment is more common among women. Nearly half of couples have both partners with a tertiary education. Educational distribution closely aligns with first birth events. Around 10% of couples include at least one partner enrolled in education, with most being single-enrollee couples. Married and cohabiting couples are equally represented, but married couples are nearly three times more likely to have children. Most couples consist of UK-born partners, and the majority of observations involve women aged 25–34, a group often partnered but childless. The summary statistics by couples' education are presented in Table A2 in the appendix.

Table 1 Sample summary statistics by the occurrence of a first birth event. Monthly observations

	Full Sample N = 87,481		First Birth Events N = 1,044	
	N	%	N	%
Live First Birth				
Yes	1,044	1.2%	1,044	100%
No	86,437	98.8%	-	-
Couples' Employment				
Both full-time	61,787	70.6%	780	74.7%
Woman part-time, Man full-time	7,936	9.1%	108	10.3%
Woman employed, Man part-time	3,369	3.9%	26	2.5%
Woman employed, Man not employed	3,482	4.0%	26	2.5%
Woman not employed, Man employed	8,317	9.5%	85	8.1%
Both not employed	2,590	3.0%	19	1.8%
Couples' Education				
Both tertiary	39,749	45.4%	485	46.5%
Heterogamous	28,937	33.1%	340	32.6%
Men higher	18,973	21.6%	246	23.5%
Women higher	9,964	11.3%	94	9.0%
Both secondary	18,795	21.5%	219	21.0%

Table 1 (continuation)

	Full Sample N = 87,481		First Birth Events N = 1,044	
	N	%	N	%
Enrolled in Education				
Not enrolled	78,705	89.6%	925	88.6%
Enrolled	9,098	10.4%	119	11.4%
Marriage Status				
Cohabitating	42,812	48.9%	267	25.6%
Married	44,669	51.1%	777	74.4%
Woman's Age				
16–24	12,163	13.9%	111	10.6%
25–29	28,773	32.9%	354	33.9%
30–34	22,859	26.1%	391	37.5%
35–39	13,229	15.1%	165	15.8%
40–45	10,457	12.0%	23	2.2%
Couples' Birth Origin				
Both UK-born	69,372	79.3%	819	78.4%
Mixed	12,163	13.9%	127	12.2%
Both foreign-born	5,946	6.8%	94	9.4%

Note: Counts represent the number of monthly observations by category.
Source: United Kingdom Household Longitudinal Study (2009–2023).

3.3 Analytical Strategy

The likelihood of the first birth was modelled using logistic regression, with standard errors clustered at the couple level, using the cross-sectional weights provided by the data managers, calculated using the *fixest* package in R (*Bergé*, 2024). Logistic regression models are considered well suited to modeling transitions to first birth in panel data (*Van Hook – Altman*, 2013). In Model 2, I introduced an interaction term between couples' employment and couples' education. The results are presented as average marginal probabilities (AMP) in Table 2. The AMPs were calculated with 95.0% confidence intervals using the *marginaleffects* package in R (*Arel-Bundock*, 2024). The AMPs of the interaction term are presented graphically for ease of interpretation following the advice of *Mize* (2019). The statistical significance of the individual predictor variables and the interaction term was evaluated through an Analysis of Deviance Type III test in Table 3.

AMPs are derived from the logistic regression models. Predicted probabilities of a first birth are calculated for each observation based on the estimated model coefficients, holding observed covariate values constant, and then averaged across the sample or relevant subgroups. AMPs therefore represent the predicted likelihood of experiencing a first birth in a given month.

While logistic regression models are estimated in odds ratios, odds ratios are difficult to interpret substantively and are not directly comparable across groups or model specifications, particularly in the presence of interaction terms. Presenting results on the probability scale facilitates interpretation and allows clearer comparison of differences in first birth likelihood across couples' employment arrangements and educational groups. Following recommended best practices for presenting nonlinear interaction effects (e.g., *Mize*, 2019), AMPs are reported in the main text, while odds ratios are provided in the Appendix for completeness (Table A3). The same set of regression models underlies all reported results; differences between tables reflect alternative presentations of identical model estimates rather than differences in model specification.

4. RESULTS

This section presents results from logistic regression models predicting the transition to first birth. It focuses on differences in average marginal probabilities across couples' employment arrangements and educational groups. All results are adjusted for woman's age, partnership status, and country of birth.

I begin by examining baseline differences in first birth probabilities across couples' employment arrangements, net of sociodemographic controls. Table 2, Model 1 presents logistic regression results for the monthly probability of first births. Couples with both partners in full-time work (AMP = 0.013, 95% CI [0.012, 0.014]) are as likely to become parents as those with full-time men and part-time women (AMP = 0.013, 95% CI [0.011, 0.016]), providing no support for H1a. In contrast, when men work part-time and women are employed, the probability is significantly lower (AMP = 0.008, 95% CI [0.005, 0.010]). The regression coefficients indicate that men's part-time employment is associated with a larger reduction in the probability of parenthood than women's part-time employment, consistent with H1b. Confidence intervals are reported to show the precision of the estimates, but they are not used to directly compare the magnitude of the effects.

Table 2 Results of logistic regression models of the monthly probability of first birth. Average marginal probabilities

First Birth Event	Model 1		Model 2	
	Estimate	CI 95%	Estimate	CI 95%
Couples' Employment				
Both full-time	0.013	(0.012, 0.014)		
Woman part-time, Man full-time	0.013	(0.011, 0.016)		
Woman employed, Man part-time	0.008	(0.005, 0.010)		
Woman employed, Man not employed	0.008	(0.005, 0.012)		
Woman not employed, Man employed	0.010	(0.008, 0.012)		
Both not employed	0.007	(0.004, 0.010)		
Couples' Education				
Both tertiary	0.012	(0.010, 0.013)		
Heterogamy	0.012	(0.011, 0.013)		
Both secondary	0.013	(0.011, 0.015)		
Enrolled in Further Education				
Not enrolled	0.012	(0.011, 0.012)	0.012	(0.011, 0.012)
Enrolled	0.013	(0.011, 0.015)	0.013	(0.011, 0.015)
Marriage Status				
Cohabitating	0.006	(0.005, 0.007)	0.006	(0.005, 0.007)
Married	0.018	(0.017, 0.020)	0.018	(0.017, 0.020)
Woman's Age				
16–24	0.014	(0.011, 0.016)	0.013	(0.011, 0.016)
25–29	0.013	(0.012, 0.015)	0.013	(0.012, 0.015)
30–34	0.016	(0.014, 0.017)	0.016	(0.014, 0.017)
35–39	0.011	(0.009, 0.012)	0.011	(0.009, 0.012)
40–45	0.002	(0.001, 0.003)	0.002	(0.001, 0.003)

Table 2				(continuation)
First Birth Event	**Model 1**		**Model 2**	
	Estimate	**CI 95%**	**Estimate**	**CI 95%**
Couples' Birth Origin				
Both UK-born	0.012	(0.011, 0.013)	0.012	(0.011, 0.013)
Mixed	0.009	(0.008, 0.011)	0.009	(0.008, 0.011)
Both foreign-born	0.014	(0.011, 0.017)	0.014	(0.011, 0.017)
Interaction Terms				
Both full-time x Both tertiary			0.013	(0.011, 0.014)
Woman part-time, Man full-time x Both tertiary			0.011	(0.007, 0.014)
Woman employed, Man part-time x Both tertiary			0.008	(0.004, 0.013)
Woman employed, Man not employed x Both tertiary			0.006	(0.002, 0.010)
Woman not employed, Man employed x Both tertiary			0.009	(0.006, 0.012)
Both not employed x Both tertiary			0.000	(0.000, 0.011)
Both full-time x Heterogamy			0.013	(0.011, 0.014)
Woman part-time, Man full-time × Heterogamy			0.014	(0.009, 0.019)
Woman employed, Man part-time × Heterogamy			0.007	(0.002, 0.012)
Woman employed, Man not employed × Heterogamy			0.008	(0.003, 0.014)
Woman not employed, Man employed × Heterogamy			0.008	(0.005, 0.011)
Both not employed × Heterogamy			0.007	(–0.001, 0.016)
Both full-time x Both secondary			0.012	(0.010, 0.014)
Woman part-time, Man full-time × Both secondary			0.017	(0.012, 0.023)
Woman employed, Man part-time × Both secondary			0.007	(0.002, 0.012)
Woman employed, Man not employed × Both secondary			0.014	(0.005, 0.022)
Woman not employed, Man employed × Both secondary			0.015	(0.009, 0.020)
Both not employed × Both secondary			0.008	(0.004, 0.013)
Observations	87,481		87,481	
AIC	10,882.4		10,889.7	
BIC	11,032.4		11,133.5	

Note: Standard errors clustered at the couple level in parentheses. Observations are measured as one observation per couple-month. Couples' characteristics are measured 12 months before the birth event. Results are average marginal probabilities derived from the same logistic regression models reported in Table A3.
Source: United Kingdom Household Longitudinal Study (2009–2023).

Table 2, Model 1 indicates that not being employed is associated with lower predicted probabilities of transitioning to parenthood: AMP = 0.008 (95% CI [0.005, 0.012]) when the man is unemployed, AMP = 0.010 (95% CI [0.008, 0.012]) when the woman is unemployed, and AMP = 0.007 (95% CI [0.004, 0.010]) when both partners are unemployed. These patterns are consistent with H1c, though the confidence intervals are provided to illustrate the precision of the estimates rather than to directly confirm the direction of the effect. This is likely due the small number of women reporting as "inactive" biasing the estimate for non-employed women downward.

In Model 1 (Table 2), couples' education shows no significant differences in first birth likelihood. Heterogamous couples and those where both have tertiary degrees have similar probabilities (AMP = 0.012) with overlapping CIs. Couples with secondary or lower education are also comparable (AMP = 0.013, 95% CI [0.011, 0.015]), supporting H2. This is reinforced by global significance tests (Table 3). However, sensitivity analysis (Table A1)

Table 3 Analysis of Deviance (Type III tests)

Analysis of Deviance	Chi-square	Df	Pr(>Chisq)
Couples' Employment	14.784	5	0.011*
Couples' Education	0.346	2	0.841
Enrolled in Further Education	3.108	1	< 0.001***
Marriage Status	254.526	1	0.078+
Woman's Age	191.165	4	< 0.001***
Couples' Birth Origin	10.847	2	0.004**
Couples' Employment x Couples' Education	12.711	10	0.240

+ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Note: Chi-square values represent likelihood ratio statistics. Df = degrees of freedom.

suggests potential differences within heterogamous couples. Those where women are higher educated show a higher likelihood of parenthood (OR = 1.169, p = 0.050) than when men are higher educated (OR = 0.809, p = 0.064), though results are marginally significant.

I next assess whether the association between couples' employment arrangements and first birth transitions differs by educational attainment by introducing interaction terms between employment status and education. Model 2 (Table 2) includes an interaction between couples' employment and education. Analysis of deviance tests (Table 3) show the interaction is not statistically significant, indicating no meaningful difference in the employment–first birth relationship across education groups.

Figure 1 visualizes the average marginal probabili ties for the interaction term with 95% confidence intervals, showing how the relationship between first birth and couples' employment varies by education. In all education groups, dual full-time and dual-earner couples with part-time working women generally show higher parenthood probabilities. Notably, the positive association between women's part-time work and parenthood appears stronger at lower education levels. However, wide confidence intervals limit broad conclusions, making within-group patterns more informative.

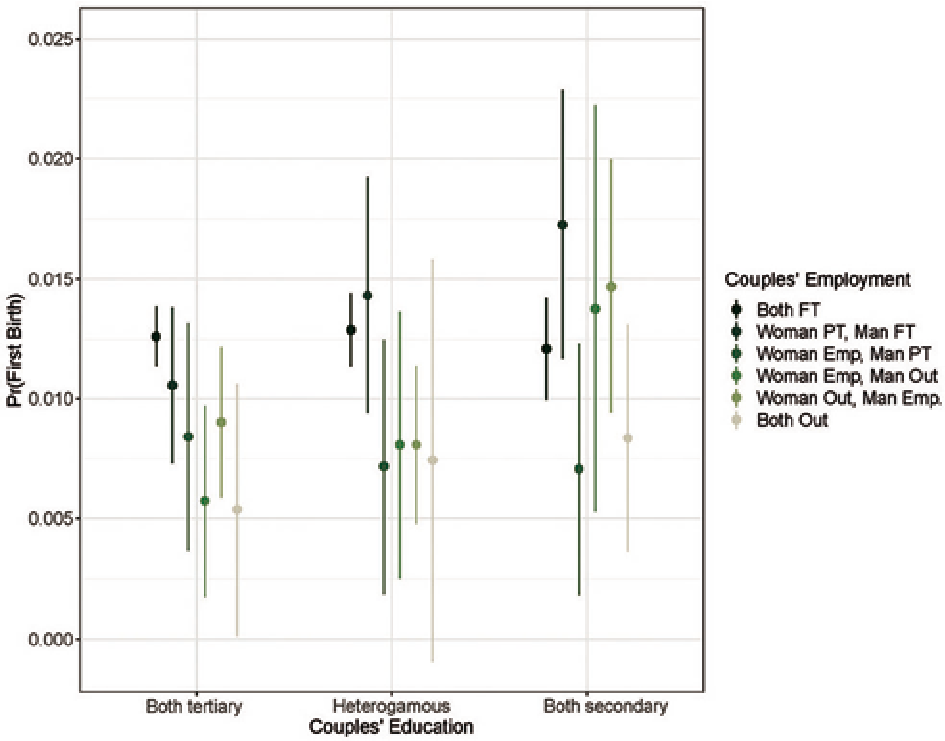
Exploring within-group variation, the relationship between employment and first birth is similar for both tertiary and heterogamous couples. In both groups, dual-earner couples with part-time working women have probabilities comparable to dual full-time earners. This probability is slightly higher for heterogamous couples (AMP = 0.014, 95% CI [0.009, 0.019]) than for couples with both partners tertiary educated (AMP = 0.011, 95% CI [0.007, 0.014], Table 2). Dual-earner couples with men in part-time work show lower probabilities, with wide confidence intervals. Single-earner couples consistently have lower probabilities, especially among heterogamous couples, where CIs are broader than in both-tertiary couples. While there's a tendency for dual employment to support parenthood among the highly educated, wide CIs limit strong conclusions, offering no clear support for H3a.

Among couples with secondary education, employment status shows a distinct pattern. Dual-earner couples with women working part-time have the highest probability of first birth—higher than dual full-time couples (AMP = 0.017, 95% CI [0.012, 0.023], Table 2)—providing tentative support for H3b. In contrast, dual-earner couples with men working part-time have the lowest probability (AMP = 0.007, 95% CI [0.002, 0.012], Table 2).

Finally, single-earner couples with non-employed women have slightly lower probabilities of first birth than dual full-time couples. However, broader confidence intervals for low-educated single-earner couples limit clear conclusions about full-time work's effect on parenthood in this group. Many missing observations from women reporting inactivity likely contribute to these wide intervals.

To conclude, although education is associated with overall levels of first birth probability, the interaction terms indicate that differences between employment arrangements are broadly similar across educational groups, suggesting limited educational moderation.

Figure 1 Predicted monthly probability of first birth by the interaction of couples' employment and couples' education



Note: This figure corresponds to Table 2, Model 2. FT = full-time employment, PT = part-time employment, Emp. = employed (full- or part-time), Out = Not employed. The model controls for enrollment in education, marriage status, couples' birth origin, and the woman's age.
Source: United Kingdom Household Longitudinal Study (2009–2023).

5. DISCUSSION

The findings align with studies from other countries showing that dual-earner couples tend to have an advantage in the transition to parenthood (*Bueno – García-Román*, 2021; *Comolli*, 2021), though this varies by working hours. While previous research suggests higher parenthood rates when both partners work full-time, this study finds similar probabilities when women work part-time and men work full-time. This may reflect the appeal of part-time work for women seeking to balance employment and caregiving or broader gendered patterns in work and family roles (*Lyonette*, 2015), though financial pressures may also play a role. This study finds some evidence that women's part-time work might be a strategy for family formation among the less educated.

Dual-earner couples with men in part-time work are less likely to become parents than those where both partners work full-time or where women work part-time and men work full-time. Their probabilities resemble those of single-earner households, suggesting that which partner holds part-time status matters. This study contributes to prior findings on part-time work by highlighting the importance of partner roles.

Unlike prior research (*Comolli*, 2021), this study does not find that single-earner couples with non-employed women are less likely to become parents than those with non-employed men. However, the interaction between employment and education reveals a complex pattern. Wide confidence intervals limit conclusions about whether women's or men's non-employment is more negatively associated with parenthood. Overall, dual-earner couples appear to have an advantage among the highly educated, but this advantage diminishes with lower educational attainment.

These findings align with broader debates on the persistence of the male breadwinner norm, highlighting that part-time work continues to be more strongly associated with delayed parenthood for men than for women. The results also underscore the value of a couple-level perspective, showing that employment arrangements cannot be fully understood in isolation but interact within partnerships to shape fertility behavior. While this study focuses on empirical patterns, these patterns suggest that gendered labor roles and social expectations continue to influence first-birth transitions in contemporary Britain.

This study has limitations. Its findings are not causal, and the logistic regression does not control for time-constant unobserved heterogeneity. The progression of education, partnership, employment, and parenthood is often nonlinear, and limited data on the type of education hinders deeper analysis. Unobserved time-varying factors like unpaid domestic labor or gender attitudes may influence outcomes, but were unavailable for monthly observations due to data limitations.

Finally, focusing on couples, while essential for understanding partners' employment dynamics, overlooks the earlier stage of partnership formation. Delays in parenthood are often linked to delays and instability in forming partnerships (*Jalovaara*, 2012). Poor labor market outcomes, especially for men, can reduce marriage prospects and thus the likelihood of parenthood (*Kalmijn*, 2011). Consequently, this study reflects a subset of the population already successful in forming partnerships.

Despite limitations, this study suggests a shift in the relationship between women's employment and parenthood in the UK since the 2008–09 financial crisis. Dual employment may be associated with facilitating parenthood, especially given rising living, childcare, and housing costs that challenge single-earner households. Among lower socioeconomic couples, women's part-time work might support parenthood, while among higher socioeconomic couples, women tend to become mothers regardless of working hours within dual-earner couples. The study finds little evidence that high-earning men encourage highly educated women to leave the workforce. Instead, for highly educated women, factors beyond economics, such as further education and career goals, may be linked to parenthood. As women's education rises, this trend will likely continue.

Acknowledgment

I would like to thank Eva Beaujouan for her dedication, invaluable insight, and guidance with this study. Isabella Buber-Ennser, Ross Barker, and Maria Rita Testa for their support and reviewers for their helpful feedback.

This study was conducted within the BIRTHLIFE project funded by the Austrian Research Foundation (FWF, P 31357-G29). A previous version of this paper was published in the Vienna Institute of Demography's Working Paper series https://epub.oeaw.ac.at/?arp=0x003d6deb.

References

- Alderotti, G. – Vignoli, D. – Baccini, M. – Matysiak, A. 2021. Employment Instability and Fertility in Europe: A Meta-Analysis. *Demography*, 58(3), 871–900. https://doi.org/10.1215/00703370-9164737.
- Arel-Bundock, V. 2024. marginaleffects: Predictions, Comparisons, Slopes, Marginal Means, and Hypothesis Tests. *R Package Version 0.21.0.* https://doi.org/10.32614/CRAN.package.marginaleffects.
- Becker, G. S. 1960. An economic analysis of fertility. In *Demographic and economic change in developed countries*, pp. 209–240. Columbia University Press. https://www.nber.org/system/files/chapters/c2387/c2387.pdf.
- Begall, K. 2013. How do educational and occupational resources relate to the timing of family formation? A couple analysis of the Netherlands. *Demographic Research*, 29, 907–936. https://doi.org/0.4054/DemRes.2013.29.34.

- Benzeval, M. – Bollinger, C. R. – Burton, J. – Crossley, T. F. – Lynn, P. 2020. *The representativeness of Understanding Society*. University of Essex. https://www.understandingsociety.ac.uk/sites/default/files/downloads/working-papers/2020-08.pdf.
- Bergé, L. 2024. fixest: Fast Fixed-Effects Estimations. *R package version 0.11.2.* https://cran.r-project.org/package=fixest.
- Blossfeld, H.-P. – Drobnic, S. 2001. *Careers of couples in contemporary society: From male breadwinner to dual-earner families: From male breadwinner to dual-earner families.* Oxford University Press.
- Budig, M. J. – Misra, J. – Boeckmann, I. 2012. The motherhood penalty in cross-national perspective: The importance of work–family policies and cultural attitudes. S*ocial Politics*, 19(2), 163–193.
- Bueno, X. – García-Román, J. 2021. Rethinking Couples' Fertility in Spain: Do Partners' Relative Education, Employment, and Job Stability Matter? *European Sociological Review*, 37(4), 571–587. https://doi.org/10.1093/esr/jcaa070.
- Cam, S. 2012. Involuntary part-time workers in Britain: Evidence from the labour force survey. *Industrial Relations Journal*, 43(3), 242–259. https://doi.org/10.1111/j.1468-2338.2012.00672.x.
- Champion, T. 2022. The United Kingdom's 'going away to university' migration and its changing impact on local populations in the twenty-first century. *Local Economy*, 37(4), 279–296. https://doi.org/10.1177/02690942221114625.
- Clark, G. – Cummins, N. 2023. Hypergamy Revisited: Marriage in England, 1837–2021. *Discussion Papers on Economics*, Article 3/2023. https://ideas.repec.org//p/hhs/sdueko/2023_003.html.
- Comolli, C. L. 2021. Couples' paid work, state-level unemployment, and first births in the United States. *Demographic Research*, 45(38), 1149–1184. https://doi.org/10.4054/DemRes.2021.45.38.
- Di Nallo, A. – Lipps, O. 2023. How much his or her job loss influences fertility: A couple approach. *Journal of Marriage and Family*, *n/a*(n/a), 1–25. https://doi.org/10.1111/jomf.12907.
- Doepke, M. – Hannusch, A. – Kindermann, F. – Tertilt, M. 2022. *The economics of fertility: A new era.* National Bureau of Economic Research. https://www.nber.org/system/files/working_papers/w29948/w29948.pdf.
- Dupray, A. – Pailhé, A. 2018. Does employment uncertainty particularly impact fertility of children of North African immigrants in France? A gender perspective. *Journal of Ethnic and Migration Studies*, 44(3), 401–424. https://doi.org/10.1080/1369183X.2017.1313107.
- Ermisch, J. 2021. English fertility heads south: Understanding the recent decline. *Demographic Research*, 45(29), 903–916. https://doi.org/10.4054/DemRes.2021.45.29.
- Esteve, A. – Schwartz, C. R. – Van Bavel, J. – Permanyer, I. – Klesment, M. – Garcia, J. 2016. The End of Hypergamy: Global Trends and Implications. *Population and Development Review*, 42(4), 615–625. https://doi.org/10.1111/padr.12012.
- Gerber, M. – Wittekind, A. – Grote, G. – Conway, N. – Guest, D. 2009. Generalizability of career orientations: A comparative study in Switzerland and Great Britain. *Journal of Occupational and Organizational Psychology*, 82(4), 779–801. https://doi.org/10.1348/096317909X474740.
- Grow, A. – Van Bavel, J. 2015. Assortative mating and the reversal of gender inequality in education in Europe: An agent-based model. *PloS One*, 10(6), e0127806. https://doi.org/10.1371/journal.pone.0127806.
- Harkness, S. – Popova, D. – Avram, S. 2023. *Gender differences in job mobility and pay progression in the UK*. Centre for Microsimulation and Policy Analysis. https://www.iser.essex.ac.uk/wp-content/uploads/files/working-papers/cempa/cempa4-23.pdf.
- Human Fertility Database. 2026a. *Total Fertility Rate—The United Kingdom*. Max Planck Institute for Demographic Research (Germany) and Vienna Institute of Demography (Austria).
- Human Fertility Database. 2026b. *Mean age at birth—United Kingdom*. Max Planck Institute for Demographic Research (Germany) and Vienna Institute of Demography (Austria).
- Inanc, H. 2015. Unemployment and the timing of parenthood: Implications of partnership status and partner's employment. *Demographic Research*, 32, 219–250. https://doi.org/10.4054/DemRes.2015.32.7.
- Jalovaara, M. 2012. Socio-economic resources and first-union formation in Finland, cohorts born 1969–81. *Population Studies*, 66(1), 69–85. https://doi.org/10.1080/00324728.2011.641720.
- Jalovaara, M. – Miettinen, A. 2013. Does his paycheck also matter? The socioeconomic resources of co-residential partners and entry into parenthood in Finland. *Demographic Research*, 28, 881–916. https://doi.org/10.4054/DemRes.2013.28.31.

- Joyner, K. – Peters, H. E. – Hynes, K. – Sikora, A. – Taber, J. R. – Rendall, M. S. 2012. The quality of male fertility data in major US surveys. *Demography*, 49(1), 101–124. https://doi.org/10.1007/s13524-011-0073-9.
- Kalmijn, M. 2011. The Influence of Men's Income and Employment on Marriage and Cohabitation: Testing Oppenheimer's Theory in Europe. *European Journal of Population / Revue Européenne de Démographie*, 27(3), 269–293. https://doi.org/10.1007/s10680-011-9238-x.
- Kreyenfeld, M. – Andersson, G. 2014. Socioeconomic differences in the unemployment and fertility nexus: Evidence from Denmark and Germany. *Advances in Life Course Research*, 21, 59–73. https://doi.org/10.1016/j.alcr.2014.01.007.
- Laß, I. 2020. The Effects of Non-standard Employment on the Transition to Parenthood Within Couples: A Comparison of Germany and Australia. *European Journal of Population*, 36(5), 843–874. https://doi.org/10.1007/s10680-019-09548-7.
- Lyonette, C. 2015. Part-time work, work–life balance and gender equality. *Journal of Social Welfare and Family Law*, 37(3), 321–333. https://doi.org/10.1080/09649069.2015.1081225.
- Lynn, P. 2009. Sample Design for Understanding Society. In *Understanding Society Working Paper* (Vols. 2009–01).
- Manning, A. – Petrongolo, B. 2008. The Part-Time Pay Penalty for Women in Britain. *The Economic Journal*, 118(526), F28–F51. https://doi.org/10.1111/j.1468-0297.2007.02115.x.
- Miettinen, A. – Jalovaara, M. 2020. Unemployment delays first birth but not for all. Life stage and educational differences in the effects of employment uncertainty on first births. *Advances in Life Course Research*, 43, 100320. https://doi.org/10.1016/j.alcr.2019.100320.
- Mikolai, J. – Berrington, A. – Perelli-Harris, B. 2018. The role of education in the intersection of partnership transitions and motherhood in Europe and the United States. *Demographic Research*, 39, 753–794. https://doi.org/10.4054/DemRes.2018.39.27.
- Mize, T. D. 2019. Best practices for estimating, interpreting, and presenting nonlinear interaction effects. *Sociological Science*, 6, 81–117. https://doi.org/10.15195/v6.a4.
- Ni Bhrolchain, M. – Beaujouan, É. 2012. France and Great Britain: Rising educational participation results in later births. *Population and Societies*, *495*. http://www.ined.fr/en/publications/pop_soc/bdd/publication/1621/.
- Nitsche, N. – Matysiak, A. – Van Bavel, J. – Vignoli, D. 2018. Partners' Educational Pairings and Fertility Across Europe. *Demography*, 55(4), 1195–1232. https://doi.org/10.1007/s13524-018-0681-8.
- OECD Statistics. 2024. *Annual average wage growth*. https://stats.oecd.org/Index.aspx?DataSetCode=AN_AV_WAGEGROWTH.
- Office for National Statistics. 2023. *Education, England and Wales—Office for National Statistics*. https://www.ons.gov.uk/peoplepopulationandcommunity/educationandchildcare/bulletins/educationenglandandwales/census2021.
- Oppenheimer, V. K. 1994. Women's rising employment and the future of the family in industrial societies. *Population and development review*, 293–342.
- Osiewalska, B. 2018. Partners' empowerment and fertility in ten European countries. *Demographic Research*, 38, 1495–1534. https://doi.org/0.4054/DemRes.2018.38.49.
- Pech, C. – Klainot-Hess, E. – Norris, D. 2021. Part-time by Gender, Not Choice: The Gender Gap in Involuntary Part-time Work. *Sociological Perspectives*, 64(2), 280–300. https://doi.org/10.1177/0731121420937746.
- Platt, L. – Knies, G. – Luthra, R. – Nandi, A. – Benzeval, M. 2020. Understanding Society at 10 Years. *European Sociological Review*, 36. https://doi.org/10.1093/esr/jcaa031.
- Savelieva, K. – Jokela, M. – Rotkirch, A. 2023. Reasons to Postpone Childbearing during Fertility Decline in Finland. *Marriage & Family Review*, 59(3), 253–276. https://doi.org/10.1080/01494929.2022.2083283.
- Scherer, S. – Brini, E. 2023. Employment Instability and Childbirth over the Last 20 Years in Italy. *European Journal of Population*, 39(1), 31. https://doi.org/10.1007/s10680-023-09680-5.
- Schmitt, C. 2012. Labour market integration, occupational uncertainties, and fertility choices in Germany and the UK. *Demographic Research*, 26, 253–292. https://doi.org/0.4054/DemRes.2012.26.12.
- Schneider, S. L. 2008. *The application of the ISCED-97 to the UK's educational qualifications*. 20.
- University Of Essex. 2023. *United Kingdom Household Longitudinal Study; Understanding Society: Waves 1-13, 2009-2023 and Harmonised BHPS: Waves 1-18, 1991-2009* (18th Edition) [dataset]. UK Data Service. https://doi.org/10.5255/UKDA-SN-6614-19.
- Van Hook, J. – Altman, C. E. 201). Using discrete-time event history fertility models to simulate total fertility rates and other fertility measures. *Population research and policy review*, 32(4), 585–610. https://doi.org/10.1007/s11113-013-9276-7.

- van Wijk, D. C. – de Valk, H. A. G. – Liefbroer, A. C. 2021. Temporary employment and family formation: An income or insecurity effect? *European Sociological Review*, 37(4), 641–658. https://doi.org/10.1093/esr/jcab007.
- Vegeris, S. – Vowden, K. – Bertram, C. – Davidson, R. – Durante, L. – Hudson, M. – Husain, F. – Mackinnon, K. – Smeaton, D. 2010. *Jobseekers Regime and Flexible New Deal Evaluation: A report on qualitative research findings*. Policy Studies Institute.
- Vignoli, D. – Drefahl, S. – De Santis, G. 2012. Whose job instability affects the likelihood of becoming a parent in Italy? A tale of two partners. *Demographic Research*, 26, 41–62. https://doi.org/0.4054/DemRes.2012.26.2.
- Warren, T. 2004. Working part-time: Achieving a successful 'work-life' balance?1. *The British Journal of Sociology*, 55(1), 99–122. https://doi.org/10.1111/j.1468-4446.2004.00008.x.
- Warren, T. – Lyonette, C. 2018. Good, Bad and Very Bad Part-time Jobs for Women? Re-examining the Importance of Occupational Class for Job Quality since the 'Great Recession' in Britain. *Work, Employment and Society*, 32(4), 747–767. https://doi.org/10.1177/0950017018762289.
- Warren, T. 2022. Work-time, male-breadwinning and the division of domestic labour: male part-time and full-time workers in unsettled times. Sociology, 56(1), 72–96. https://doi.org/10.1177/00380385211019660.
- Wright, L. 2020. *Producing working-life histories in the BHPS and UKHLS 2017-2020* [dataset]. UK Data Service. https://doi.org/10.5255/UKDA-SN-854327.

BRIAN BUH

A demographer who received his doctorate in Demography from the University of Vienna. His research focuses on fertility behaviour in low-fertility contexts, with particular attention to the interplay between childbearing and competing life domains such as employment, housing, and partnership trajectories. He has worked as a Research Assistant at the Vienna Institute of Demography, Austrian Academy of Sciences, contributing to projects including *BIRTHLIFE – Competing Life Course and Fertility* and *KriLebWie – Fertility in Times of Crises*. He is the author and co-author of several peer-reviewed publications on fertility, residential mobility, employment uncertainty, and environmental behaviour, published in journals such as *Population Studies*, *Genus*, and *Travel Behaviour and Society*.

APPENDIX

Table A1 Sensitivity analysis of explanatory variables couples' employment and couples' education. Monthly probability of first birth. Model 1 disaggregates women's non-employment between not employed and carework/homemaker. Model 2 disaggregates heterogamous education by which partner has a higher level of education. Odds ratios.

First Birth Event	Model (1)	Model (2)
Couples' Employment		
Ref. Both full-time	*Ref.*	–
Woman part-time, Man employed	0.983 (0.100)	–
Woman employed, Man not employed	0.661* (0.133)	–
Woman not employed, Man employed	0.620** (0.097)	–
Woman carework/homemaker, Man employed	1.167 (0.193)	–
Both not employed	0.566* (0.134)	–
Couples' Education		
Ref. Both tertiary	*Ref.*	–
Heterogamy	1.033 (0.074)	–
Both secondary	1.104 (0.095)	–
Couples' Employment		
Ref. Both full-time	–	Ref.
Woman part-time, Man full-time	–	1.042 (0.109)
Woman employed, Man part-time	–	0.599* (0.120)
Woman employed, Man not employed	–	0.646* (0.130)
Woman not employed, Man employed	–	0.802+ (0.094)
Both not employed	–	0.553* (0.131)
Couples' Education (4 categories)		
Ref. Both tertiary	–	Ref.
Woman tertiary, Man secondary	–	1.169+ (0.093)
Woman secondary, Man tertiary	–	0.809+ (0.092)
Both secondary	–	1.126 (0.096)
Enrolled in Further Education		
Ref. Not enrolled	*Ref.*	*Ref.*
Enrolled	1.134+ (0.078)	1.130+ (0.077)
Marriage Status		
Ref. Cohabitating	*Ref.*	*Ref.*
Married	3.109*** (0.237)	3.159*** (0.241)
Woman's Age		
Ref. 25–29	*Ref.*	*Ref.*
16–24	1.027 (0.117)	1.024 (0.117)
30–34	1.161* (0.087)	1.156+ (0.087)
35–39	0.780** (0.075)	0.788* (0.076)
40–45	0.140*** (0.030)	0.139*** (0.030)

Table A1		(contunuation)
First Birth Event	**Model (1)**	**Model (2)**
Couples' Birth Origin		
Ref. Both UK-born	*Ref.*	*Ref.*
Mixed	0.763** (0.075)	0.764** (0.075)
Both foreign-born	1.109 (0.124)	1.148 (0.128)
Intercept	0.007*** (0.001)	0.007*** (0.001)
Observations	87,481	87,481
AIC	10,882.6	10,874.9
BIC	11,032.6	11,034.4

+ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Note: Standard errors clustered at the couple level in parentheses. Observations are measured as one observation per couple-month. Couples' characteristics are measured 12 months before the birth event. Results are based on logistic regression models predicting the transition to first birth. Standard errors clustered at the couple level in parentheses. Observations are measured as one observation per couple-month. Couples' characteristics are measured 12 months before the birth event. Results are based on logistic regression models predicting the transition to first birth.
Source: United Kingdom Household Longitudinal Study (2009–2023).

Model 1 shows that couples with women engaged in care work or homemaking tend to have higher – but not statistically significant – likelihoods of becoming parents, consistent with prior UK research (*Inanc, 2015; Schmitt, 2012*). However, many "inactive" women were excluded, limiting conclusions about inactivity and parenthood.

Model 2 reveals that accounting for couples' education changes the likelihood of parenthood for heterogamous couples; specifically, couples where women are more educated than their partners show a positive association with first birth.

Table A2 Summary Statistics for the Sample by Couples' Education. Monthly Observations

	Both tertiary N=39,749	Heterogamous N=28,937	Both secondary N=18,795
Live first birth			
Yes	1.2%	1.2%	1.2%
No	98.8%	98.8%	98.8%
Couples' employment status			
Both full-time	73.1%	74.7%	59.1%
Woman part-time, Man full-time	9.3%	7.2%	11.6%
Woman employed, Man part-time	3.8%	3.5%	5.1%
Woman employed, Man not employed	3.8%	3.5%	4.6%
Woman not employed, Man employed	8.2%	9.4%	12.5%
Both not employed	2.0%	1.7%	7.1%
Enrolled in Education			
Not enrolled	90.9%	85.4%	91.9%
Enrolled	9.1%	14.6%	8.9%
Marriage status			
Cohabitating	44.8%	50.7%	55.0%
Married	55.2%	49.3%	45.0%
Woman's age			
16–24	8.2%	12.7%	27.9%
25–29	35.1%	32.5%	28.8%
30–34	28.6%	26.4%	20.4%
35–39	16.3%	15.8%	11.7%
40–45	11.9%	12.5%	11.2%
Couples' birth origin			
Both UK-born	74.5%	82.0%	85.2%
Mixed	16.1%	12.8%	10.9%
Both foreign-born	9.4%	5.2%	3.9%

Note: Counts represent the number of monthly observations by category.
Source: United Kingdom Household Longitudinal Study (2009–2023).

Table A3 Results of logistic regression models of the monthly probability of first birth. Odds ratio

First Birth Event	Model (1)	Model (2)
Couples' Employment		
Ref. Both full-time	*Ref.*	*Ref.*
Woman part-time, Man full-time	1.043 (0.109)	0.831 (0.139)
Woman employed, Man part-time	0.593** (0.119)	0.666 (0.196)
Woman employed, Man not employed	0.701* (0.130)	0.448* (0.161)
Woman not employed, Man employed	0.779* (0.091)	0.708+ (0.132)
Both not employed	0.547* (0.130)	0.423+ (0.214)
Couples' Education		
Ref. Both tertiary	*Ref.*	*Ref.*
Heterogamy	1.042 (0.075)	1.023 (0.083)
Both secondary	1.115 (0.096)	0.949 (0.101)
Enrolled in Further Education		
Ref. Not enrolled	*Ref.*	*Ref.*
Enrolled	1.128+ (0.077)	1.134+ (0.078)
Marriage Status		
Ref. Cohabitating	*Ref.*	*Ref.*
Married	3.190*** (0.244)	3.193*** (0.244)
Woman's Age		
Ref. 25–29	*Ref.*	*Ref.*
16–24	0.855 (0.115)	0.840 (0.114)
30–34	1.442** (0.168)	1.441** (0.168)
35–39	1.290 (0.292)	1.295 (0.292)
40–45	0.321** (0.130)	0.322** (0.130)
Couples' Birth Origin		
Ref. Both UK-born	*Ref.*	*Ref.*
Mixed	0.772** (0.075)	0.773** (0.076)
Both foreign-born	1.152 (0.128)	1.160 (0.129)
Interaction Terms		
Ref. Both full-time x Both tertiary		*Ref.*
Woman part-time, Man full-time × Heterogamy		1.330 (0.336)
Woman employed, Man part-time × Heterogamy		0.83 (0.403)
Woman employed, Man not employed × Heterogamy		1.377 (0.701)
Woman not employed, Man employed × Heterogamy		0.866 (0.248)
Both not employed × Heterogamy		1.349 (1.041)
Woman part-time, Man full-time × Both secondary		1.734* (0.441)
Woman employed, Man part-time × Both secondary		0.858 (0.421)
Woman employed, Man not employed × Both secondary		2.522+ (1.232)

Table A3 (continuation)

First Birth Event	Model (1)	Model (2)
Woman not employed, Man employed × Both secondary		1.697+ (0.471)
Both not employed × Both secondary		1.588 (0.936)
Intercept	0.012*** (0.003)	0.013*** (0.003)
Observations	87,481	87,481
AIC	10,878.4	10,885.7
BIC	11,037.9	11,138.9

+ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Note: Standard errors clustered at the couple level in parentheses. Observations are measured as one observation per couple-month. Couples' characteristics are measured 12 months before the birth event. Odds ratios are based on the same regression models as Table 2 and are presented for completeness.
Source: United Kingdom Household Longitudinal Study (2009–2023).

HOUSEHOLD ECONOMIC STATUS AND MIGRATION: EVIDENCE FROM CENTRAL AND EASTERN EUROPEAN COUNTRIES

Oana-Maria Cozma[1)] – **Andrei-Ionuț Pricop**[2)]

Abstract
Migration has been a defining phenomenon in Eastern Europe since the fall of the Iron Curtain and has intensified with the free movement of people within the European Union. This paper examines the relationship between migration and household economic conditions, an area that has received limited empirical attention. Using panel data for 11 Central and Eastern European EU member states from 2013 to 2023, the study analyses how household financial indicators influence net migration balances. The empirical analysis applies panel regression techniques, including pooled OLS, random effects, and fixed effects models. The results show that household net financial assets and household investment rates are significant predictors of migration outcomes, while household savings and debt exhibit weaker or insignificant effects. These findings suggest that household-level economic resources and investment capacity play an important role in shaping migration dynamics in Eastern Europe. By connecting standardized household economic indicators to migration patterns, the paper provides new comparative evidence with relevant implications for migration and economic policy in the region.

Keywords: Household Economic Indicators, Migration, Panel Data Analysis, Eastern European Countries, Financial Assets and Investment
Demografie, 2026, **68(1): 25–35**
DOI: https://doi.org/10.54694/dem.0377

INTRODUCTION

Migration has been a defining socioeconomic phenomenon in Eastern Europe since the early 1990s, when the fall of the Iron Curtain opened new opportunities for mobility (*Favell*, 2008). The process intensified after EU enlargement, as the principle of free movement facilitated large-scale emigration from Eastern to Western Europe (*Kahanec – Zimmermann*, 2016). From a theoretical perspective, these migration flows are commonly explained by international migration theories that emphasize economic differentials between sending and receiving countries, such as neoclassical economic theory and the new economics of labour migration, which highlight wage gaps, employment opportunities, and household income strategies as key drivers of mobility (*Todaro*, 1969; *Harris – Todaro*, 1970; *Stark – Bloom*, 1985). This dynamic has profoundly shaped labour markets, household structures, and national

1) Faculty of Economics and Business Administration, Alexandru Ioan Cuza University of Iași, Romania. Contact: cozma.oanamaria@gmail.com.
2) Faculty of Economics and Business Administration, Alexandru Ioan Cuza University of Iași, Romania. Contact: ap.pricop@gmail.com.

development strategies, making it a central topic for policymakers and researchers (*OECD*, 2019). Eastern European countries were selected because they have experienced some of the most intense migration dynamics in Europe since the 1990s, combining large-scale emigration, rapid economic transformation, and significant changes in household financial structures. This makes them particularly relevant for analyzing how household economic indicators influence migration, in ways that may not be as visible in more stable Western economies. The reason for this study comes from the recognition that, while migration flows are widely analyzed from macroeconomic and labour market perspectives, the role of household-level financial indicators remains underexplored. Existing research has mainly focused on how migration affects household outcomes, including remittances, savings, and investments (*Bauer – Sinning*, 2005; *Incalțărău – Maha*, 2012). While these studies demonstrate that migration reshapes household financial behavior, much less is known about the reverse mechanism, whether and how household financial conditions themselves drive or reduce migration. Standardized indicators such as the gross household saving rate, household debt-to-income ratio, household net financial assets, and household investment rates remain largely absent from comparative empirical analyses (*Lou*, 2022; *Seitz*, 2019).

The purpose of this paper is twofold: first, to provide a targeted literature review that identifies what is already known and what remains underexplored in this field; and second, to empirically test how household financial indicators shape migration balances across Eastern European EU countries. These aims are important because household-level financial resilience or vulnerability may influence migration decisions, linking microeconomic behavior with macro-level demographic outcomes (*McKenzie – Rapoport*, 2007; *Gentili*, 2013). The rationale of the present paper lies in bridging microeconomic household indicators with migration dynamics at the country level. By applying panel data econometric methods (OLS, random effects, and fixed effects with Hausman tests) to 11 Eastern European EU states between 2013 and 2023, the paper offers systematic comparative evidence on the relationship between household saving, debt, assets, investment, and net migration balances (*Wooldridge*, 2010). Developing this idea further, the paper contributes to a deeper understanding of migration as not only a labour market phenomenon but also a reflection of household financial conditions. By distinguishing the relative importance of assets, savings, debt, and investments, and by comparing high- and moderate-migration countries, this research provides new insights into the differentiated mechanisms through which household financial structures shape migration in Eastern Europe (*Dossche et al.*, 2022; *Kozlovskyi et al.*, 2024).

LITERATURE REVIEW ON HOUSEHOLD ECONOMIC INDICATORS AND NET MIGRATION DYNAMICS

The present literature review examines studies that research the relationship between household economic indicators and net migration dynamics. A targeted literature review was conducted in two major academic databases, Web of Science and Scopus, both recognized for their extensive coverage of peer-reviewed research. The search strategy relied on specific keyword combinations aimed at capturing the relationship of household economic indicators and migration, using semantic variations and Boolean operators. These included: "household economic indicators AND migration", "household saving rates AND migration", "household net financial assets AND migration", and "household investment rates AND migration". The results revealed that only a limited number of papers directly address this specific nexus, highlighting a gap in the existing literature (Table 1 summarizes the literature review).

Research on the relationship between household economic indicators (HEIs) and migration dynamics reveals diverse mechanisms depending on geography, level of analysis, and methodological approach. A first distinction must be made between studies that examine how migration reshapes household indicators, and those that test whether household conditions themselves drive migration decisions. Several papers focus on the impact of migration on household savings and investment behavior. *Bauer and Sinning* (2005), using German GSOEP data, show that immigrants generally save less than natives, but that this gap disappears for temporary migrants once remittances are considered; return intentions significantly increase both remittances and savings. Their findings underline how migration intentions shape household financial strategies rather than the reverse. Similarly, *Incalțărău and Maha* (2012), analyzing Romanian macro data, conclude that remittances contributed more to household and aggregate investment than to consumption, highlighting migration as both a survival and capital-formation strategy. *Lou* (2022) extends this discussion to rural China, finding that household migration is negatively associated with saving rates, while other controls such as education, health, marital status, and insurance also reduce savings. These studies show the migration and household indicators channel but do not establish whether strong household conditions affect migration itself (*Bauer – Sinning*, 2005; *Încalțărău – Maha*, 2012; *Lou*, 2022).

By contrast, evidence from Uzbekistan points to a clearer push mechanism. Using national surveys, *Seitz* (2019) finds that households with low life satisfaction, unemployment, or weak labor market integration are more likely to consider migration, while actual migrant households enjoy higher income and improved well-being. Remittances play a central role in reducing poverty (from 16.8% to 9.6% in counterfactual scenarios), smoothing shocks and improving life satisfaction. Here, poor household conditions clearly enhance migration, while improvements in income and satisfaction would be expected to reduce outflows (*Seitz*, 2019).

Theoretical and empirical work on migration as a household financial strategy further sharpens this link. *Clemens and Ogden* (2014) conceptualize migration as an investment decision under household credit and insurance constraints, with remittances as the returns. They argue that migration is chosen when local financial instruments are unavailable or inadequate, implying that improvements in household financial stability would reduce migration. Mexican evidence is particularly illustrative. *McKenzie and Rapoport* (2007) and *Gentili* (2013) both find an inverted U-shaped relationship between household wealth and migration. Poor households face liquidity constraints and cannot finance migration; middle-wealth households migrate most; and wealthier households reduce migration as local opportunities increase. Migration networks lower costs over time, allowing poorer households to participate. Thus, "better" HEIs initially enhance migration up to a threshold, but at higher wealth levels they reduce migration (*McKenzie – Rapoport*, 2007; *Gentili*, 2013).

Studies of internal migration in advanced economies provide complementary insights. *Olney and Thompson* (2024), analyzing U.S. IRS data show that internal migration decreases with higher origin wages and increases with higher destination wages. Rising housing costs historically spurred migration, but households have become less responsive to these costs over time, contributing to the long-run decline in mobility. This suggests that strong household conditions (higher wages, affordable housing) reduce the need for migration, while weak conditions (low wages, high housing costs) enhance it (*Olney – Thompson*, 2024).

At a macro level, *Dossche et. al.* (2022) documents that immigrants in the euro area have systematically lower wages and wealth, are more liquidity-constrained, and are more sensitive to labor market cycles, while the immigrant share has increased by 5 percentage points since 2007. *Kozlovskyi et al.* (2024), using Eurostat data, show positive correlations between net migration and GDP in many EU countries, with particularly strong effects in Poland, Spain, and the Baltics. These studies underline how migration changes the distribution of household economic indicators and macroeconomic performance, though they do not directly measure how household conditions affect migration rates (*ECB*, 2022; *Kozlovskyi et al.*, 2024).

Taken together, the literature suggests that the relationship between household economic indicators and migration is context-dependent and often nonlinear. Evidence from developing countries (Uzbekistan, rural Mexico, rural China) highlights how poor household indicators (low income, low satisfaction, weak financial tools) drive migration as a survival or portfolio strategy. Yet, once households accumulate sufficient wealth, the need to migrate diminishes. In advanced economies, higher wages and affordable housing anchor populations, while deteriorating conditions push migration. At the aggregate European level, net migration inflows are associated with stronger economic performance, showing how migration in turn reshapes household outcomes. Thus, the balance of evidence suggests that good household economic indicators generally reduce migration at higher levels of welfare, but at low and middle levels, improvements in income or wealth may initially facilitate migration by relaxing financial constraints.

RESEARCH GAPS

The targeted literature review shows that while the relationship between migration, remittances, and household financial behavior has been widely studied, the direct role of household economic indicators such as the gross household saving rate (HSR), household debt-to-income ratio (HDEB), household net financial assets ratio (HNFR), and the gross investment rate of households (HIR) remains largely unexplored. Existing studies mostly analyze how migration reshapes savings and investment, rather than whether these indicators influence migration decisions. Evidence is also fragmented by geography, concentrated in single-country cases outside the EU, with limited cross-country comparative work using harmonized data. Moreover, the literature rarely

Table 1 Literature review on household economic indicators (HEIs) and migration

Authors & year	Geographical coverage	Relevance for household economic indicators and migration dynamics	Do "better" household economic indicators reduce or enhance migration?
Lou (2022)	Rural China	The direction studied is migration to savings (not savings to migration). Shows how migration changes HEIs (savings behaviour).	Not estimated. No evidence that higher savings/other HEIs raise or lower migration rates.
Bauer & Sinning (2005)	Germany	Links migration intent/ duration to household saving/ remitting behaviour and composition of migration (temporary vs permanent) shapes HEIs.	Not estimated. This paper doesn't test whether higher savings/wealth change migration rates.
Dossche et. al. (2022)	Euro area	Documents systematic HEI gaps by migrant status and notes rising immigration; i.e., migration reshapes the distribution of HEIs and macro transmission.	Not estimated. Descriptive on HEIs and migrant shares; no causal effect of HEIs on migration rates.
Încalţărău & Maha (2012)	Romania	Shows migration - remittances -household/ aggregate investment channel (HEIs respond to migration).	Not estimated. No test of whether household investment/consumption levels change migration rates.
Seitz (2019)	Uzbekistan	Clear push mechanism from poor HEIs to migration; migration then improves HEIs via income/ remittances.	Better HEIs reduce out-migration: when jobs/ income/ life satisfaction improve, the push weakens. Evidence is strongest on the push side.
Clemens & Ogden (2014)	Global (conceptual)	Theoretical link: household financial constraints and risk management are core drivers of migration dynamics.	Conceptual: easing constraints / improving HEIs would reduce the need for migration as a risk-management tool.
McKenzie & Rapoport (2007)	Mexico	Direct micro-mechanism tying household wealth and net emigration dynamics (via networks & costs).	Non-monotonic: better HEIs (wealth) enhance migration from low to middle, but reduce it at high wealth.
Kozlovskyi et al. (2024)	Europe (EU-27, CH, UA)	Macro evidence of migration affecting development; does not unpack household-level mechanisms.	Not estimated. Paper relates net migration to GDP, not HEIs - migration.
Gentili (2013)	Mexico	Strong household-finance angle: wealth, assets, networks jointly determine migration intensity.	Non-monotonic: better HEIs (wealth) enhance migration up to mid-wealth, then reduce it at higher wealth.
Olney & Thompson (2024)	United States (internal)	Clean price/ income mechanisms: household income & housing costs shape internal migration patterns and their decline.	Higher origin wages (good HEI) reduce out-migration; higher destination wages enhance migration; higher origin housing costs push out-migration (but responsiveness has fallen).

Source: Authors' elaboration.

addresses dynamics over time or disentangles causality between household conditions and migration flows. These gaps suggest the need for systematic empirical research linking household-level and macro-level indicators directly to migration patterns, particularly in the Eastern EU context.

DATA AND METHODS

This study investigates the relationship between household economic indicators and net migration dynamics in Eastern European countries through a panel data analysis. The analysis covers 11 EU member states from Eastern Europe (Bulgaria, Croatia, Czechia, Estonia, Latvia, Lithuania, Poland, Romania, Slovakia, Slovenia, and Hungary) over the period 2013–2023. The dependent variable is the crude rate of net migration plus statistical adjustment, as reported by Eurostat. Net migration was chosen because the balance between emigration and immigration captures not only emigration decisions but also the migration flows that have shaped these countries in recent years. In particular, during the COVID-19 pandemic, Poland experienced a sharply negative migration balance (−20.3 in 2020, followed by 0.1 in 2021), which subsequently recovered, a pattern not observed in the other countries in the sample. By contrast, the invasion of Ukraine, which began on 24 February 2022, generated an exceptionally positive migration balance for most of the countries studied, especially those in close geographic proximity to Ukraine (with net migration increasing from 4.8 in 2021 to 30.9 in 2022 in the Czech Republic; from 5.3 to 29.5 in Estonia; from −0.2 to 11.7 in Latvia; from 7.0 to 25.5 in Lithuania; and from −0.9 to 5.6 in Romania).

The independent variables include four household economic indicators: (i) Gross Household Saving Rate (HSR), (ii) Household Debt-to-Income Ratio (HDEB), (iii) Household Net Financial Assets Ratio (HNFR), and (iv) Gross Investment Rate of Households (HIR). These variables capture different dimensions of household financial capacity and resilience.

Table 2 shows the descriptive statistics for the variables used in the analysis.

It can be observed that the descriptive statistics reflects different patterns and distributions for the analyzed variables.

The analysis began with pooled Ordinary Least Squares (OLS) regressions to establish baseline associations, followed by Random Effects (RE) and Fixed Effects (FE) estimations to account for unobserved country-specific heterogeneity. The Hausman specification test was applied to choose between RE and FE models. Results from the Hausman test generally favored the FE specification, which was adopted as the preferred model for interpretation. The baseline empirical model can be expressed as:

$$\mathbf{MIGRATION}_{it} = \alpha + \beta 1\mathbf{HSR}_{it} + \beta 2\mathbf{HNFR}_{it} + \beta 3\mathbf{HIR}_{it} + u,$$

where MIGRATION$_{it}$ represents the net migration rate in country i at time t, while the coefficients β capture the marginal effects of household economic indicators. Panel unit root tests were not performed because, in small panels such as the one used in this study, these tests are not a necessary condition and may, moreover, distort the overall statistical analysis (*Mravak*, 2023; *Jönsson*, 2005). Nevertheless, the possibility of conducting stationarity tests in future analyses with a longer time dimension is not excluded.

Table 2 Descriptive statistics

Indicator	HDEBT	HIR	HNFR	HSR	MIGRATION
Mean	47.49	7.30	135.48	7.21	0.66
Median	46.06	7.17	133.61	8.55	0.25
Maximum	72.56	14.42	255.53	22.88	30.90
Minimum	21.81	1.17	64.22	−11.18	−20.30
Std. Dev.	13.22	2.40	40.65	7.18	6.74
Observations	120	120	120	120	120

Source: Authors' estimations.

RESULTS, LIMITATIONS AND DISCUSSION

Results

Given its consistently low statistical significance, HDEB was excluded in the second model to improve estimation efficiency, while HSR was also excluded in Model 3. As shown in Table 3, the final preferred model (Model 3, FE) includes HNFR and HIR.

To further capture potential heterogeneity in migration dynamics, the sample was further divided into two groups: (1) High-emigration countries, characterized by persistently negative net migration balances (Romania, Bulgaria, Croatia, Latvia, Lithuania) and (2) Moderate-emigration countries, with migration rates closer to balance or fluctuating around zero (Czechia, Slovenia, Slovakia, Estonia, Hungary, Poland). As observed in Table 3 and Table 4, separate FE and RE estimations for each sub-sample reveal distinct patterns: in high-emigration countries, HNFR emerges as the key predictor of migration, whereas in moderate-emigration countries, HIR plays the central role.

Table 3 Baseline regression results for household economic indicators and net migration (2013–2023)

	Model 1 OLS	Model 1 Random Effects	Model 1 Fixed Effects	Model 2 OLS	Model 2 Random Effects	Model 2 Fixed Effects	Model 3 OLS	Model 3 Random Effects	Model 3 Fixed Effects
HSR	0.141 (0.087)	0.080 (0.109)	−0.054 (0.152)	0.149 (0.079) *	0.092 (0.101)	−0.053 (0.149)	–	–	–
HNFR	0.046 (0.014) ***	0.072 (0.020) ***	0.142 (0.036) ***	0.045 (0.014) ***	0.068 (0.018) ***	0.143 (0.035) ***	0.055 (0.013) ***	0.077 (0.018) ***	0.139 (0.033) ***
HDEB	0.010 (0.044)	0.007 (0.061)	−0.007 (0.122)	–	–	–	–	–	–
HIR	1.258 (0.225) ***	1.467 (0.305) ***	1.346 (0.536) **	1.258 (0.224) ***	1.444 (0.291) ***	1.352 (0.525) **	1.253 (0.226) ***	1.468 (0.304) ***	1.356 (0.522) **
C	−16.305 (3.497) ***	−20.751 (4.903) ***	−27.792 (9.075) ***	−15.759 (2.689) ***	−19.788 (3.355) ***	−28.248 (4.504) ***	−15.960 (2.716) ***	−20.570 (3.468) ***	−28.165 (4.480) ***
R^2	0.280	0.253	0.451	0.280	0.253	0.451	0.258	0.249	0.450
Hausman	Chi sq. 8.830			Chi sq. 9.452			Chi sq. 7.328		
	Prob. 0.0655			Prob. 0.0238			Prob. 0.0256		

Note: Significance levels: * for >10%, ** for > 5%, *** for >1%.
Source: Authors' estimations.

Table 4 Regression results with only high-emigration countries

	Random Effects Model 1	Fixed Effects Model 1	Random Effects Model 2	Fixed Effects Model 2
HSR	0.202 (0.153)	0.091 (0.189)	–	–
HNFR	0.111 (0.029) ***	0.208 (0.048) ***	0.115 (0.028) ***	0.231 (0.043) ***
HIR	1.702 (0.416) ***	0.912 (0.645)	1.606 (0.403) ***	0.883 (0.637)
C	−27.891 (5.481) ***	−34.412 (5.912) ***	−27.354 (5.409) ***	−34.870 (5.788) ***
R^2	0.311	0.502	0.287	0.499
Hausman	Chi-Sq. 10.548		Chi-Sq. 11.565	
	Prob. 0.0144		Prob. 0.0030	

Note: Significance levels: * for >10%, ** for > 5%, *** for >1%.
Source: Authors' estimations.

Table 5 Regression results with only moderate-emigration countries

	Random Effects Model 1	Fixed Effects Model 1	Random Effects Model 2	Fixed Effects Model 2
HSR	–0.177 (0.204)	–0.293 (0.228)	–	–
HNFR	0.052 (0.030) *	0.058 (0.053)	0.038 (0.023) *	0.041 (0.051)
HIR	1.662 (0.554) ***	2.474 (0.851) ***	1.553 (0.493) ***	2.385 (0.853) ***
C	–15.684 (4.555) ***	–21.499 (6.527) ***	–14.984 (4.036) ***	–21.726 (6.560) ***
R^2	0.235	0.387	0.237	0.370
Hausman	Chi-Sq. 3.019		Chi-Sq. 2.513	
	Prob. 0.3886		Prob. 0.284	

Note: Significance levels: * for >10%, ** for > 5%, *** for >1%.
Source: Authors' estimations.

Panel data models are particularly suited for this analysis because they allow for controlling unobserved heterogeneity across countries, exploit both cross-sectional and time-series variation, and increase estimation efficiency compared to pure cross-sectional or time-series approaches (*Wooldridge*, 2010).

For high-emigration countries, the fixed effects model is more appropriate, as the Hausman test yields p-values below 0.05. By contrast, for countries with moderate emigration, a random effects model is preferred, and the results are therefore less conclusive.

A linear model was chosen because it directly captures household well-being in relation to emigration decisions and destination attractiveness for immigrants. Moreover, given the size of the sample, a linear specification is considered more appropriate than nonlinear models (*Wooldridge*, 2023).

The baseline regressions for the full sample of 11 Eastern European countries (2013–2023) reveal that household net financial assets (HNFR) and the gross investment rate of households (HIR) are consistently significant predictors of net migration balances, while the gross household saving rate (HSR) shows weaker and inconsistent effects. The household debt-to-income ratio (HDEB) proved statistically insignificant across all specifications and was therefore excluded in the preferred model (Model 2, FE). Model 2 with fixed effects, supported by the Hausman test, indicates that higher household net financial assets and higher household investment rates are associated with more positive migration balances, suggesting that periods of improved household wealth and investment correspond to reduced emigration or increased immigration in the region. By contrast, saving rates did not exhibit a robust relationship with migration.

The sub-sample analysis uncovers important heterogeneity. In high-migration countries, HNFR is the main driver, highlighting the role of household wealth and capital accumulation in shaping migration decisions. Conversely, in moderate-migration countries, HIR becomes more relevant, suggesting that local household investments are more effective in anchoring populations when migration pressures are less acute. These findings underline the differentiated mechanisms through which household financial conditions influence migration depending on national migration contexts.

Contribution to the literature

This study addresses the identified gaps by providing direct empirical evidence on how household economic indicators shape migration in Eastern European EU member states. Using panel data for 2013–2023 and testing four household-level indicators (HSR, HDEB, HNFR, HIR), the results show that household net financial assets (HNFR) and household investment rates (HIR) are significant predictors of net migration balances, while household debt (HDEB) and savings rates (HSR) play only marginal roles. Importantly, the analysis distinguishes between high-migration and moderate-migration countries: in the former, household assets matter more, while in the latter, investment rates are the main driver. By linking standardized Eurostat indicators to migration flows through fixed-effects models, this research fills the measurement, regional, and temporal gaps in the literature, offering new comparative evidence that highlights the differentiated mechanisms through which household financial conditions influence migration dynamics in Eastern Europe.

As control variables, the analysis includes three indicators from the Heritage Economic Freedom Index (2024), namely Labor Freedom, Monetary Freedom, and Property Rights. Table 6 reports the results starting from the baseline model, in which Property Rights emerges as the only statistically significant control variable:

The results reported in Table 6 indicate the importance of property rights in migration decisions.

Next, Pesaran's CD test (*Pesaran*, 2004) and the Wooldridge test (*Wooldridge*, 2002) are applied to assess issues related to cross-sectional dependence and autocorrelation.

Table 6 Regressions with control variables

	Model 1 Random Effects	Model 1 Fixed Effects	Model 2 Random Effects	Model 2 Fixed effects	Model 3 Random Effects	Model 3 Fixed Effects
HSR	0.017 (0.081)	−0.146 (0.149)	0.022 (0.076)	−0.149 (0.148)	0.025 (0.077)	−0.110 (0.145)
HNFR	0.041 (0.014) ***	0.126 (0.035) ***	0.039 (0.013) ***	0.126 (0.035) ***	0.038 (0.013) ***	0.119 (0.035) ***
HIR	0.715 (0.249) ***	0.689 (0.558)	0.698 (0.232) ***	0.706 (0.555)	0.714 (0.234) ***	0.681 (0.556)
LABOR FREEDOM	−0.003 (0.057)	−0.053 (0.085)	–	–	–	-
PROPERTY RIGHTS	0.195 (0.038) ***	0.138 (0.048) ***	0.198 (0.037) ***	0.137 (0.048) ***	0.197 (0.037) ***	0.142 (0.048) ***
MONETARY FREEDOM	−0.051 (0.143)	−0.179 (0.154)	−0.042 (0.139)	−0.185 (0.154)	–	-
C	−19.001 (11.586)	−11.881 (13.676)	−19.792 (11.366) *	−14.728 (12.844)	−23.150 (2.815) ***	−29.231 (4.364) ***
R^2	0.402	0.501	0.414	0.499	0.409	0.492
Hausman	Chi-Sq. 10.685		Chi-Sq. 10.493		Chi-Sq. 7.750	
	Prob. 0.098		Prob. 0.062		Prob. 0.101	

Note: Significance levels: * for >10%, ** for > 5%, *** for >1%.
Source: Authors' estimations.

Table 7 Pesaran CD test and Wooldridge test for Model 3 with fixed effects from Table 3

Wooldridge	F (1, 10) = 0.108	Prob > F = 0.7495
Pesaran CD-Test	Pesaran's test of cross sectional independence = 5.885, Pr = 0.0000	Average absolute value of the off-diagonal elements = 0.310

Source: Authors' estimations.

Table 8 Regression with Drisc/Kraay errors

HNFR	0.139 (0.041) ***
HIR	1.356 (0.702) *
C	–28.165 (6.841) ***
R^2	0.281
Obs.	120

Note: Significance levels: * for >10%, ** for > 5%, *** for >1%.
Source: Authors' estimations.

The test results suggest re-estimating the model using Driscoll–Kraay standard errors. Table 8 reports the regression results for Model 3 from Table 3 estimated with Driscoll–Kraay standard errors.

Limitations

Several limitations should be acknowledged. First, the analysis relies on aggregate Eurostat indicators, which may obscure household-level heterogeneity and within-country regional disparities. Second, the study does not fully address potential endogeneity between migration and household indicators; while higher wealth may reduce emigration, migration itself can shape wealth accumulation through remittances and labour market effects. Third, the focus on Eastern EU member states ensures comparability but limits generalizability to non-EU contexts or Western Europe. Finally, the ten-year window (2013–2023) captures important post-crisis dynamics but may not fully reflect longer-term structural transformations in household finance and migration behavior.

Discussion

The results provide new evidence on the underexplored nexus between household financial conditions and migration dynamics. The strong effects of household net financial assets and investment rates highlight that migration is not solely a labour market phenomenon but also reflects the financial resilience of households. The fact that HNFR matters most in high-migration contexts suggests that capital constraints remain a key driver of emigration, consistent with theories of migration as a household investment decision (*Clemens – Ogden*, 2014; *McKenzie – Rapoport*, 2007). In moderate-migration contexts, the relevance of household investments indicates that local opportunities and capital formation can anchor populations more effectively, in line with development-driven migration theories. The insignificance of debt and the weak role of saving rates are also noteworthy. While debt dynamics often dominate household finance discussions, they appear less relevant for migration decisions in Eastern Europe, perhaps because credit markets remain less developed compared to Western economies. Similarly, saving rates may reflect precautionary behaviour rather than structural financial capacity, limiting their explanatory power. Taken together, these findings advance the literature by linking standardized household indicators directly to migration outcomes and by showing that their effects are context-dependent. They also carry implications for policy, strengthening household asset accumulation and facilitating productive investment may reduce emigration pressures, especially in high-migration countries, while in moderate-migration contexts, supporting household-level investment opportunities could stabilize migration balances.

CONCLUSION

This paper has examined the relationship between household economic indicators and migration in Eastern European EU countries over the period 2013–2023. By using panel data econometric methods, the study has demonstrated that household net financial assets and household investment rates are significant predictors of net migration balances, while household saving rates play only a marginal role and household debt levels appear irrelevant. The results highlight important heterogeneity across contexts. In high-migration countries, household assets are more influential, reflecting the economic vulnerabilities and capital needs that sustain emigration pressures. In moderate-migration countries, household investments matter more, pointing to the role of local opportunities in anchoring populations when migration dynamics are less acute. These findings contribute to filling a notable gap in the literature, which has focused predominantly on remittances and post-migration household outcomes rather than pre-migration financial conditions. By linking standardized Eurostat household indicators directly to migration flows, this paper advances both the empirical and conceptual understanding of migration as a phenomenon shaped not only by labour market dynamics but also by household financial structures. At the same time, the analysis underlines that the effects of household economic conditions on migration are neither linear nor uniform. Instead, they are context-dependent and interact with broader economic development patterns. Policymakers in Eastern Europe may thus reduce emigration pressures by supporting household wealth accumulation and facilitating productive investments, while recognizing that the financial drivers of migration differ across high- and moderate-migration countries. Future research could build on this study by incorporating micro-level household data, addressing endogeneity concerns, and extending the analysis beyond Eastern Europe to assess whether similar patterns hold in other regional and institutional contexts.

References

- Bauer, T. K. – Sinning, M. 2005. The savings behavior of temporary and permanent migrants in Germany, *IZA Discussion Paper No. 1632*. IZA – Institute of Labor Economics. https://doi.org/10.2139/ssrn.742330.
- Clemens, M. – Ogden, T. N. 2014. Migration as a strategy for household finance: A research agenda on remittances, payments, and development, *Working Paper No. 354*. Center for Global Development. https://doi.org/10.2139/ssrn.2457148.
- Dossche, M. – Kolndrekaj, A. – Propst, M. – Ramos Perez, J. – Slacalek, J. 2022. Immigrants and the distribution of income and wealth in the euro area: First facts and implications for monetary policy, *ECB Working Paper Series No. 2719*. European Central Bank. https://doi.org/10.2139/ssrn.4219401.
- Favell, A. 2008. *Eurostars and eurocities: Free movement and mobility in an integrating Europe* (1st ed.). Wiley. https://doi.org/10.1002/9780470712818.
- Gentili. 2013. Migration Costs and Networks: household optimal investment in migration, *Quaderni - Working Paper DSE, No. 867*, Alma Mater Studiorum Università di Bologna, Dipartimento di Scienze Economiche (DSE), Bologna. https://doi.org/10.6092/unibo/amsacta/3896.
- Harris, J. R. – Todaro, M. P. 1970. Migration, unemployment and development: A two-sector analysis. *American Economic Review*, 60(1), 126–142.
- Încălțărău, C. – Maha, L.-G. 2012. The impact of remittances on consumption and investment in Romania. *Eastern Journal of European Studies*, 3(2), 61–73.
- Jönsson, K. 2005. *Testing for Stationarity in Panel Data when Errors are Serially Correlated. Finite-Sample Results.* Working Papers 16. Lund University, Department of Economics.
- Kahanec, M. – Zimmermann, K. F. (Eds.). 2016. *Labor migration, EU enlargement, and the great recession*. Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-662-45320-9.
- Kozlovskyi, S. – Kulinich, T. – Vechirko, I. – Lavrov, R. – Zayukov, I. – Mazur, H. 2024. Relationship between net migration and economic development of European countries: Empirical conclusions. *Problems and Perspectives in Management*, 22(1), 605–618. https://doi.org/10.21511/ppm.22(1).2024.48.
- Lou, F. 2022. Examining the relations between household saving rate of rural areas and migration. *arXiv*. https://doi.org/10.48550/ARXIV.2201.07159.
- Mckenzie, D. – Rapoport, H. 2007. Network effects and the dynamics of migration and inequality: Theory and evidence from Mexico. *Journal of Development Economics*, 84(1), 1–24. https://doi.org/10.1016/j.jdeveco.2006.11.003.

- Mravak, I. 2023. A Monte Carlo study on the size and power of panel unit root tests: Limitations in small data sets. *Croatian Operational Research Review*, 14(2). https://doi.org/10.17535/crorr.2023.0011.
- OECD. 2019. *Talent Abroad: A Review of Romanian Emigrants*, OECD Publishing, Paris. https://doi.org/10.1787/bac53150-en.
- Olney, W. – Thompson, O. 2024. The determinants of declining internal migration. *National Bureau of Economic Research*. https://doi.org/10.3386/w32123.
- Pesaran, M. H. 2004. General diagnostic tests for cross section dependence in panels. *CESifo Working Paper Series No. 1229*. https://doi.org/10.2139/ssrn.572504.
- Seitz, W. 2019. International migration and household well-being: Evidence from Uzbekistan. *Policy Research Working Paper No. 8910*. World Bank Group. https://doi.org/10.1596/1813-9450-8910.
- Stark, O. – Bloom, D. E. 1985. The new economics of labor migration. *American Economic Review*, 75(2), 173–178.
- The Heritage Foundation 2024. *2024 Index of Economic Freedom*, Washington, DC 20002. The Heritage Foundation.
- Todaro, M. P. 1969. A model of labor migration and urban unemployment in less developed countries. *American Economic Review,* 59(1), 138–148.
- Wooldridge, J. M. 2002. *Econometric Analysis of Cross Section and Panel Data*. MIT Press.
- Wooldridge, J. M. 2023. Simple approaches to nonlinear difference-in-differences with panel data. *The Econometrics Journal*, 26(3). https://doi.org/10.1093/ectj/utad016.

OANA-MARIA COZMA

is an Associated Teaching Assistant (PhD) in the Economics and International Relations Department at the Faculty of Economics and Business Administration, 'Alexandru Ioan Cuza' University of Iași, Romania. Her research interests include labour force migration, with a focus on Romanian labour migration within the European Union, as well as slavery and the history of economic thought.

IONUȚ-ANDREI PRICOP

is PhD candidate and Associated Teaching Assistant in the Economics and International Relations Department at "Alexandru Ioan Cuza" University of Iași, Romania. His research interests are economic security, economic freedom, human capital and human development, international economics and European studies.

Využití cenzového informačního systému pro populační statistiku

Zdroje dat pro oficiální populační statistiku[1)] v Česku se postupně odklání od tradičních demografických hlášení a sčítání zaměřeného primárně na sčítací formuláře a směřuje ke sběru dat zejména pomocí administrativních zdrojů dat (AZD), mezi kterými má nejpodstatnější úlohu cenzový informační systém (dále jen CENZIS). CENZIS byl zřízen v návaznosti na Zákon o sčítání lidu, domů a bytů v roce 2021[2)] a je ukotven od srpna 2020 v Zákoně o státní statistické službě[3)]. Jedná se o informační systém, jehož správcem je Český statistický úřad (dále jen ČSÚ), a který má sloužit ke splnění povinností stanovených přímo použitelnými předpisy Evropské unie v oblasti migrace a mezinárodní ochrany, evropské demografické statistiky, sčítání lidu, domů a bytů. V této oblasti statistiky došlo v listopadu 2025 ke sjednocení do jednotného evropského nařízení „European Statistics on Population and Housing“ (ESOP)[4)].

OBSAH CENZISu

V současnosti lze demograficky relevantní obsah CENZISu rozdělit do tří skupiny: data ze sčítání 2021, registr obyvatel a jeho agendové informační systémy a ostatní administrativní zdroje dat. Zákon o sčítání lidu 2021 umožnil uchovat v CENZISu individuální data o obvyklém pobytu v rozhodný okamžik, obvyklém pobytu v době narození a nejvyšším ukončeném vzdělání včetně nezbytných identifikátorů osob pro další využití v populační statistice.

CENZIS dále tvoří průběžně aktualizovaná soustava registrů ROB–AISEO–AISC[5)]. Aktualizace probíhají v ROBu na základě notifikací, v průběhu roku 2026 je plánován přechod na systém notifikací i u agendových informačních systémů. Časová prodleva mezi časem demografické události a její evidencí v CENZISu je tak dána pouze dobou zápisu do informačních systémů veřejné správy. CENZIS přebírá z ROB–AISEO–AISC pouze položky relevantní pro populační statistiku, jde například o základní identifikaci osoby, rodinný stav / (registrované) partnerství, datum narození a případně úmrtí, adresu registrovaného pobytu, státní občanství, obec / stát místa narození a případně úmrtí, typ pobytu, vazbu na rodinné příslušníky, některé historické údaje a mnohé další.

Mezi ostatní AZD patří databáze/evidence dalších rezortů, které jsou importovány do CENZISu s roční, nebo půlroční periodicitou. Jedná se o Centrální registr pojištěnců spravovaný VZP (zahrnuje všechny zdravotní pojišťovny a eviduje všechny osoby registrované k veřejnému zdravotnímu pojištění). Dále pak informační systémy MPSV – OK Práce (obsahuje data o uchazečích a zájemcích o zaměstnání), OK Nouze (společně posuzované osoby u doplatku na obydlení, příspěvku na péči a příspěvku na živobytí) a OK Centrum (společně posuzované osoby u přídavku na dítě, příspěvku na bydlení a rodičovského příspěvku). Česká správa sociálního zabezpečení poskytuje ze svého Integrovaného informačního systému subsystém důchodů včetně druhu důchodu. Od MŠMT jsou předávány údaje ze školních matrik (základní, střední, vyšší odborné školy a konzervatoře) a sdružené informace matrik studentů (vysokoškolská

1) Pro účely tohoto příspěvku se populační statistikou rozumí průběžná demografická statistika i data ze sčítání lidu, domů a bytů za osoby, rodiny a domácnosti.

2) Zákon č. 332/2020 Sb.

3) Zákon č. 89/1995 Sb.

4) Nařízení Evropského parlamentu a Rady (EU) č. 2025/2458 ze dne 26. listopadu 2025 o evropských statistikách obyvatelstva a bydlení.

5) Registr obyvatel – Agendový informační systém evidence obyvatel – Agendový informační systém cizinců.

studia). Jedná se o osoby navštěvující dané typy škol s údaji o případném ukončení studia. Posledním subjektem, který v současnosti dodává do CENZISu data je Centrum pro zjišťování výsledků vzdělávání (CZVV). Z informačního systému IS CERTIS se jedná o údaje za studenty, kteří se přihlásili k maturitní zkoušce včetně informace o jejím získání. V blízké budoucnosti se počítá s daty Jednotného měsíčního hlášení zaměstnavatelů zejména pro ekonomické charakteristiky.

VYUŽITÍ CENZISu

CENZIS je zapojen do sběru a zpracování dat pro běžnou evidenci demografických událostí od roku 2024, kdy se stal zdrojem dat pro sňatky a zemřelé osoby v případě občanů Česka. Od roku 2026 jsou nově z CENZISu generovány i narozené děti a partnerství českých občanů. Úmrtí či narození cizinců a sňatky/partnerství, kde je aspoň jeden ze snoubenců/partnerů cizinec dále podléhají statistickým hlášením u úmrtí, narození dítěte a uzavření manželství/partnerství, které vyplňují a odesílají na ČSÚ matriční úřady. Data z AISC v současnosti bohužel nesplňují kvalitativní parametry (neúplnost, velké zpoždění), aby mohla být použita pro podchycení demografických událostí cizinců. Předpokládá se, že situaci by mohl zlepšit nový Integrovaný cizinecký agendový systém (ICAS), který se v současnosti buduje.

Údaje z demografických hlášení se podařilo téměř plně nahradit s využitím dat výše uvedených informačních systémů. V některých případech to ovšem nebylo možné. Data o příčinách úmrtí stále pochází z Listu o prohlídce zemřelého. Porodní charakteristiky narozeného dítěte dostupné registry v CENZISu neobsahují, takže budou využita mikrodata Ústavu zdravotnických informací a statistiky ČR. V některých případech došlo k výraznému zlepšení vyplněnosti údajů, které se začaly přebírat z CENZISu. Například u nejvyššího ukončeného vzdělání zemřelých, rodičů narozených dětí a snoubenců.

Do budoucna se předpokládá využití CENZISu i pro účely statistiky rozvodů/ukončení (registrovaných) partnerství soudem, jež v současnosti předává ČSÚ Ministerstvo spravedlnosti. Další využití může být u evidence stěhování občanů Česka, případně i cizinců (nejdříve se zprovozněním ICASu). Výhodou proti současnému stavu (data jsou přebíraná od Ministerstva vnitra a cizinecké policie) by byla možnost sledování stěhování i v rámci obce.

Cílovým stavem u populační statistiky je ukončení klasických bilancí obyvatelstva vázaných na data sčítání lidu a přechod na populační stavy odvozené z dat CENZISu. K tomu je potřeba harmonizovat údaje o stavu a pohybu obyvatel zpracované z CENZISu zejména v návaznosti na „signs of life“ analýzu (dále jen SOL), která umožní identifikovat osoby, které na území Česka již nežijí. Ke kvalitnímu provedení SOL je nutné využití širokého aparátu administrativních zdrojů dat, které CENZIS obsahuje. Přechod na odvození populačních struktur z CENZISu je nutný i kvůli ESOPu, který rozšiřuje množství povinně předávaných dat s nižší než desetiletou periodicitou na údaje, které je v současnosti možné získat pouze ze sčítání lidu (například populační struktury podle ekonomické aktivity). Vzhledem k tomu, že sčítání lidu není možné realizovat vícekrát během dekády, tak je nutné využít AZD, které jsou, nebo budou v budoucnu dostupné v CENZISu. Zároveň se samozřejmě počítá s využitím CENZISu pro příští sčítání 2031/2032, respektive pro témata, která je nutné poskytovat podle nařízení ESOP v desetileté periodicitě.

PŘÍKLAD VYUŽITÍ CENZISU – ODVOZENÍ NEJVYŠŠÍHO UKONČENÉHO VZDĚLÁNÍ

Od roku 2005 je uvedení údaje o nejvyšším ukončeném vzdělání na demografických hlášeních dobrovolné. V návaznosti na tuto úpravu začala vyplněnost proměnné vzdělání klesat až na hodnoty, kdy údaj ztrácí vypovídací schopnost. Například v roce 2023 nebyl v případě zemřelých údaj o vzdělání vyplněn v 84 % případů. Prostředí CENZISu umožňuje odhadnout nejvyšší ukončené vzdělání pro populaci evidovanou v registru obyvatel. K tomu bylo využito množství různých datových zdrojů – data ze sčítání 2021, školní matriky a sdružené informace matriky studentů od MŠMT[6)], údaje o maturitních

6) Kromě informací o probíhajících a ukončených studií obsahují i údaj o nejvyšším ukončeném vzdělání před nástupem na daný typ studia.

zkouškách od CZVV a evidenci uchazečů a zájemců o zaměstnání[7] od MPSV. Výsledné odvozené nejvyšší ukončené vzdělání je nutné vnímat jako určitý odhad. Využívá různorodé zdroje dat, pro které má údaj o vzdělání odlišnou relevanci. Zároveň se nejedná o registr, který by byl průběžně aktualizován, ale spíše o odhad k začátku roku na základě nejnovější dostupných AZD, které pokrývají téměř celou populaci[8]. Po napojení na demografická hlášení tak nemusí být odhadnuta úroveň vzdělání, které osoba měla přesně v době demografické události. Na druhou stranu průběžně probíhající analýzy kvality těchto dat zatím potvrdily, že odhady jsou relevantní. Od roku 2024 vzniká nejvyšší ukončené vzdělání u zemřelých osob, snoubenců a rodičů narozených dítě jako kombinace vyplněných údajů na hlášeních a odhadnuté úrovně vzdělání z CENZISu, přičemž údaje z hlášení mají prioritu. Tabulka 1 porovnává struktury zemřelých podle vzdělání v roce 2004 (poslední rok, kdy bylo vyplnění na hlášení povinné), 2023 (poslední rok s daty pouze z hlášení) a 2024 (první rok s kombinovaných způsobem). Po zapojení CENZISu v roce 2024 klesl podíl nezjištěných údajů o vzdělání z cca 84 % na zhruba 9 %. Zároveň je patrné, že při srovnání s rokem 2004 vzrostla vzdělanostní úroveň zemřelých díky tomu, že se do věků s vyšší mírou úmrtnosti postupně posouvají vzdělanější generace. Nízký podíl neúplných dat o vzdělání byl v roce 2004 evidován i v případě sňatků (7 % u nevěst a 8 % v případě ženichů) a narozených dětí (2 % u matek a 3 % v případě otců).

ZÁVĚR

Zapojení administrativní zdrojů dat do tvorby oficiální populační statistiky ČSÚ se v kontextu mezinárodních statistických trendů zdůrazňující digitalizaci, snížení zátěže pro poskytovatele dat a právě využívání AZD stalo nutností, které přináší mnoho výhod. Data mohou být zpracována rychleji, efektivněji, přesněji, v elektronickém formátu a s nižšími nároky na rozpočet. Důležitá je i možnost propojování různých

Tab. 1: Zemřelí podle nejvyššího ukončeného vzdělání v letech 2004, 2023 a 2024 (v %)
Deaths by educational attainment, 2004, 2023 and 2024 (in %)

Nejvyšší dokončené vzdělání	2004	2023	2024
Bez vzdělání	0,6	0,3	0,3
Základní	46,4	2,8	19,1
Střední bez maturity (vč. vyučení)	39,5	6,8	40,9
Střední s maturitou	10,6	3,7	22,1
Vyšší odborné	*	0,1	0,7
Vysokoškolské	3,0	2,7	8,1
Nezjištěno	0,0	83,6	8,7
Celkem	100,0	100,0	100,0

Pozn.: * vyšší odborné vzdělání nebylo rozlišeno.
Note: ** Short-cycle tertiary education was not distinguished.*
Zdroj: Demografická ročenka České republiky 2004, 2023 a 2024.
Source: *Demographic Yearbook of the Czech Republic 2004, 2023 a 2024.*

7) Obsahuje informaci o nejvyšším ukončeném vzdělání.

8) Z dostupných zdrojů nemáme informaci o dosaženém vzdělání na školách v zahraničí a zahraničních pobočkách škol na českém území.

zdrojů dat. V některých případech jsou získané údaje dokonce úplnější. Neplatí to ovšem vždy. Určité proměnné není možné ze současných AZD v Česku získat (například kulturních charakteristiky, domácnosti, dojížďku), a proto úplných přechod na administrativní sčítání není zatím možný. Nevýhodou využití AZD je skutečnost, že jsou primárně používána pro jiné účely, takže musí být před jejich použitím ověřena jejich kvalita a metodický soulad s produkovanými statistikami. Problematické mohou být i omezené možnosti vzájemného propojování a v některých případech neúplné pokrytí populace. Rozvoj využití administrativních zdrojů dat v populační statistice Česka bude v následujících letech pro odbor statistiky obyvatelstva ČSÚ prioritním úkolem.

Roman Kurkin
Odbor statistiky obyvatelstva, ČSÚ

The 17th Conference of Young Demographers

On 4–6 February 2026, around 90 participants gathered for the 17th Conference of Young Demographers, which was held at the Faculty of Science, Charles University in Prague. Over the course of three days, early-career researchers from across Europe and beyond presented their work on health, mortality, fertility, inequalities, migration, and demographic theory in a welcoming and intellectually stimulating environment.

The conference opened on Wednesday morning in the Velká geologická lecture room. Session 1, chaired by Anna Altová, primarily focused on health and ageing. Florencia Bathory presented a multistate analysis of chronic conditions and depression onset among older Europeans. Andrea Cass examined the link between pre-retirement working conditions and long-term care needs among the oldest old in Germany. Maria Chiara Corgiolu explored ethnic inequalities in the use of telehealth in the United States, and Federica Prete discussed typologies of family and non-family networks in later life and their association with health changes.

The first poster session covered a broad spectrum of demographic topics: gender inequalities in outsourcing domestic work during widowhood, social media use and fertility intentions among young adults in Europe, regional heterogeneity in the effectiveness of training policies for older unemployed workers, and delayed transitions to adulthood among young people living with their parents in contemporary Europe. As is tradition, each poster session began with flash talks, in which presenters had two minutes to introduce their research before in-depth discussions continued in the poster rooms.

After lunch, Katarina Luise Matthes delivered the first keynote lecture: *'Born during the 1918–20 Pandemic: What Historical Data Reveal About Neonatal Health and Later Mortality'*. In addition to describing her interesting work, she also walked participants through her career: starting as an engineering student and becoming a historical demographer.

The afternoon concluded with a second poster session. The presentations addressed topics such as 19th-century child mortality in Sweden, the demographic ageing of physicians in Czechia, widowhood trends worldwide, the third demographic transition, vaccination and fertility patterns, and many other topics. The variety of themes reflected the breadth of contemporary demographic research.

Thursday began with a second oral session. Sofía García-Durrer examined the relationship between perceived unfairness in the division of household labour and union dissolution among German couples. Sina Christin Löwe explored how unpaid care work shapes parental well-being in Germany. Sharon Picco analysed interparental relational mismatch and its implications for children's non-cognitive development. Francesco Tata focused on the demography and social gradient of online dating in Italy. Together, the session offered a multifaceted perspective on family dynamics, gender roles, and contemporary relationship patterns.

Poster Session 3 continued the discussion of inequality and demographic change. Contributions ranged from defining the demographic crisis, suicide ideation among older adults in Europe, temperature-related mortality inequalities in Belgium, the victimisation of transgender people in Brazil, cross-border reproductive care in Europe, and lifespan inequality among migrants and natives.

Marie-Pier Bergeron Boucher delivered the second keynote lecture, titled *'Can Simple Measures Capture Complex Realities?'* She provided insight into her work on demographic indicators and measurement. However, she also shared her experience of being both a mom and a researcher, showing that it is possible to have both a family life and a successful career in science.

Session 3, chaired by Klára Hulíková Tesárková, featured four presentations focused on life-course trajectories, health, and social inequalities across Europe. Beatrice Cacciamani presented a cross-country analysis of how parental separation shapes children's educational and occupational outcomes later in life. Patrik Harnisch examined trends and heterogeneity in the use of eContacts in Denmark between 2005 and 2019, drawing on register-based data. Lina Lasar analysed health expectancy among Roma populations across European countries from a comparative perspective. Pilar Vidaurre-Teixidó concluded the session with a presentation on variations in the social determinants of health across 27 European countries in 2024, using principal component analysis to uncover cross-national patterns.

The day concluded with the Giuliana Award session, a creative and much-anticipated highlight of the conference that celebrated outstanding contributions from early-career researchers. This was followed by a convivial social evening, which provided participants with the opportunity to network, engage in informal discussions, and enjoy light refreshments in a relaxed and collegial atmosphere.

The final day opened with Session 4, chaired by Věra Slováková, and focused primarily on mortality, demographic methods, and long-term population trends. Paul-Émile Charlier investigated the deficit of daughter-only families in South Asia, proposing a measure based on fertility tables to examine prenatal discrimination behaviours. Dariya Ordanovich reconstructed the mortality burden from temperature extremes in Madrid, Spain, over the course of 130 years using age-specific mortality rates. David A. Swanson presented a regionally based probabilistic forecast of human extinction, offering a thought-provoking methodological perspective on long-term demographic forecasting. Tamara Vaz concluded the session by reassessing mortality trends in Canada, the United States, and the United Kingdom, benchmarking them against Japan, and critically discussing the limitations of life expectancy as a summary indicator.

Poster Session 4 provided the final opportunity for discussion. Topics included fertility clustering methods, the demographic impacts of COVID-19 in Portugal, low family aspirations in South Korea, the use of machine learning in regional population projections, student mobility, ageing in Australia, cohort contributions to population growth in Czechia, colorectal cancer mortality trends, and demographic divergence in the post-Yugoslav region.

The conference concluded with a closing ceremony and the announcement of awards celebrating the conference's outstanding presentations and posters. Beatrice Cacciamani received the Best Presentation Award and Riccardo Dalla Tomba won the Best Poster Award. Elder Lara Castañed received the 'Giuliana Award'. As in previous years, the defining features of the conference were its constructive feedback, supportive atmosphere, and high level of scholarly exchange.

In addition to the main programme of the 17th Conference of Young Demographers, two specialised COST GREATLEAP masterclasses, supported by the COST (European Cooperation in Science and Technology) framework, were offered to registered conference participants. Due to limited capacity and high demand, participation in the masterclasses was subject to prior registration and confirmation.

The first masterclass, held in the evening of Wednesday, 4 February 2026, was delivered by Katarina Luise Matthes and entitled *'How to Build a Strong Research Proposal – Using Examples from Research on Health Inequalities'*. The session provided structured guidance on the preparation of competitive research proposals, with a particular emphasis on formulating research questions, methodological design, and the coherent presentation of research objectives and expected contributions. Examples from research on health inequalities were used to illustrate key principles and best practices.

The second masterclass took place on Thursday, 5 February 2026, and was led by Marie-Pier Bergeron-Boucher. Entitled *Health and Mortality Measures*, the session focused on advanced measures of health and mortality and addressed their conceptual foundations, methodological construction, interpretation, and application in demographic research.

Both masterclasses, organised under the COST GREATLEAP initiative, complemented the main conference programme by offering participants an opportunity for in-depth methodological training, direct interaction with internationally recognised scholars, and further professional development in a focused academic setting. The organisers are pleased and grateful that it was possible to organise these masterclasses, which served as an exceptional opportunity to enhance the conference experience for all participants.

The organisers would like to express their sincere gratitude to the Faculty of Science, Charles University, for hosting the event and to the Department of Demography and Geodemography for its institutional support. We are especially grateful to our sponsors, whose support made the conference possible: the N. W. Posthumus Institute, the European Association of Historical Demography, SAS, and the Czech Statistical Office.

The 17th Conference of Young Demographers once again demonstrated the vitality and diversity of contemporary demographic research. The Young Demographers Team looks forward to welcoming participants again next year for the 18th conference.

The Young Demographers Team:
Anna Altová, Kateřina Brázová, Klára Hulíková Tesárková, Barbora Janáková Kuprová, Lukáš Kahoun, Jitka Slabá, and Bety Ukolova.

E-DEMOGRAPHY AS A PLATFORM FOR POPULATION DATA ANALYSIS: A CONCEPTUAL REVIEW AND THE CASE OF AZERBAIJAN

Farhad Yusifov[1)] – Elvin Gurbanov[2)]

Abstract
Electronic demography (e-demography) refers to the application of digital technologies and big data in demographic research and population-related decision-making. This study explores the concept, importance, and evolution of e-demography through a comprehensive review of the literature and a comparative assessment of international practices. Particular emphasis is placed on the role of government registers in the development of e-demography systems and the implementation of digital platforms. The study presents Azerbaijan as a practical context by briefly describing the population register and the Azerbaijan Service and Assessment Network (ASAN) Service digital platform as enabling components for e-demography-oriented applications. In addition, the study discusses how e-demography complements classical demographic infrastructure (including vital statistics) and it outlines key challenges and opportunities, such as data quality, interoperability, privacy protection, and institutional capacity. The results highlight the role of digital transformation in demographic management and underscore the potential of global best practices to inform the development of e-demography in Azerbaijan. Implications include enhancing e-government service quality, increasing citizen satisfaction, and improving oversight of socio-political processes. Future research will focus on expanding socio-demographic analyses and applying advanced analytical methods.

Keywords: Electronic demography, Population register, Government registers, Big data, Digital footprints, Artificial intelligence, Machine learning, Internet of Things, ASAN Service
Demografie, 2026, **68(1): 42–51**
DOI: https://doi.org/10.54694/dem.0375

INTRODUCTION

The rapid advancement of information and communication technologies has driven digital transformation across socio-economic domains, profoundly influencing both the field of demography and the practice of population governance. This study is structured as a review, synthesising current literature and international practices related to e-demography and digital population governance. Traditionally, demographic research relied on large-scale data sources such as censuses, civil registration records, and surveys. However, in recent decades, the digitalisation of society has generated unprecedented volumes of digital data about people's activities and the traces of those activities. In this context, the concept of electronic demography (e-demography)

1) Institute of Information Technology, B.Vahabzade str., 9A, Baku, AZ1141, Azerbaijan. Contact: farhadyusifov@gmail.com.
2) Institute of Information Technology, B.Vahabzade str., 9A, Baku, AZ1141, Azerbaijan. Contact: elvingurbanov@list.ru.

has emerged to denote the integration of digital data and technologies into the study of demographic processes and into the decision-making on population management. In this study, the term "e-demography system" refers to an integrated socio-technical arrangement in which digital data sources (for example, registers and platforms) are linked with analytical methods to support demographic monitoring and policy-oriented decision-making.

In classical demography, researchers have long utilised big datasets, censuses, and registration systems containing millions of individual records. In the digital age, the notion of big data encompasses diverse, high-velocity, real-time data streams. Demographers can now harness new-generation data sources such as users' digital footprints (for example, social media activity, mobile phone location data) (*Kashyap*, 2021). Alongside these novel data, the application of advanced analytical methods and computational technologies has given rise to a subfield often termed computational or digital demography. *Kashyap and Zagheni* (2023) note that e-demography investigates how emerging digital data streams and computational techniques contribute to understanding population dynamics, while concurrently examining the effects of digital technologies on demographic outcomes such as health, fertility, and migration. This development expands the boundaries of demographic science, enabling more timely and comprehensive information to inform policy decisions.

The importance of e-demography is underscored by modern governments' need for timely and accurate data to develop sustainable demographic policies. Trends such as population ageing, migration flows, urbanisation, and fertility rates evolve rapidly, and traditional decennial census data may prove insufficiently responsive. In this regard, integrating state information resources and establishing real-time statistical monitoring systems are essential for informed demographic policymaking. E-demography addresses this requirement by merging data from multiple government registers, digital platforms, and alternative sources, thus facilitating operational analytical assessments of the current population status, trend dynamics, and forecasts. Consequently, the e-demography paradigm has crystallised in response to the digital revolution within the demographic data ecosystem and has assumed a critical role in both academic research and governmental decision-making practice. Finally, since not all countries operate population registers, classical demographic infrastructure, particularly civil registration and vital statistics, remains a foundational component of demographic measurement and should be considered alongside newer digital data sources.

This article addresses key issues in e-demography, integrating conceptual, infrastructural, operational, and technological perspectives. Section 2 reviews conceptual frameworks and international practices. Section 3 examines population registers as data sources; Section 4 illustrates the realization of ASAN Service in Azerbaijan; Section 5 explores technological advances such as AI, machine learning, remote sensing, and IoT; and Section 6 discusses challenges and opportunities for building robust, ethical, and sustainable e-demography systems.

CONCEPTUAL FRAMEWORKS AND INTERNATIONAL PRACTICES IN E-DEMOGRAPHY SYSTEMS

Recent studies in e-demography indicate that the discipline of demography has undergone a profound transformation under the influence of the big data era (*Poulain – Herm*, 2013; *Cesare et al.*, 2018; *Alguliyev et al.*, 2019; *Yusifov*, 2021; *Kashyap – Zagheni*, 2023; *Civelek*, 2023). This section synthesises key scholarly perspectives on e-demography and reviews selected international practices in digital demographic data systems relevant to population governance. *Kashyap* (2021) observes that although demographers have traditionally been accustomed to handling large-scale datasets, the proliferation of digital data generated by a digitising society not only creates new opportunities but also poses ethical and technical challenges. The study highlights two primary shifts in the data ecosystem: first, the enhancement and increased interoperability of traditional data sources (censuses, surveys, administrative registries); and second, the incorporation of digital footprints arising from internet and mobile device usage into research (*Kashyap*, 2021). In this context, *Cesare et al.* (2018) argue that digital traces of online activities (for example, social media posts, search queries) are attractive data sources for the real-time monitoring of demographic processes. However, they caution that issues of privacy, representativeness, and data quality must be carefully addressed, noting that digital traces cover only internet users and thus may under-represent certain social groups (*Cesare et al.*, 2018). *Yusifov* (2021) proposes a conceptual model for an e-demography system in Azerbaijan that is based on public registers, emphasising the use of personal identification numbers to integrate individual-level data. Leveraging these administrative data sources can significantly enhance demographic research by enabling more detailed analyses of population characteristics. *Kashyap and Zagheni* (2023) propose integrating policy-oriented simulation models into e-demographic research. By generating "synthetic populations" through agent-based models and microsimulation techniques, they explore the potential to virtually assess the demographic impacts of social policies. Importantly, these simulation-based synthetic populations are typically used for scenario testing and require calibration/validation against empirical demographic data; they are no substitute for real-world measurement. Such approaches permit the pre-evaluation of long-term outcomes for policies aimed at, for example, stimulating population growth or strengthening social protections for the elderly (*Kashyap – Zagheni*, 2023). Additional studies support this paradigm. *Juranek and Zoutman* (2021) utilised mobile phone data to assess the effects of social restrictions during the pandemic, while *Wilde et al.* (2020) attempted to forecast US birth rates using Google search queries. Similarly, *Zagheni et al.* (2019) proposed statistical correction models to track migrant flows via social media advertising platforms (*Zagheni et al.*, 2019; *Wilde et al.*, 2020; *Juranek – Zoutman*, 2021).

Nevertheless, literature reviews reveal that demographers have not reached a full consensus regarding big data. *Wunsch et al.* (2024), in a systematic analysis, found that researchers generally prefer official and statistical sources, treating non-traditional digital data with caution. This reluctance stems partly from methodological challenges in establishing causal inferences and partly from insufficient ethical and legal regulations (*Wunsch et al.*, 2024). This debate also highlights a practical point: population registers are not universal, and many countries continue to rely primarily on civil registration and vital statistics alongside other official sources. These theoretical and methodological perspectives are reflected in the practical models adopted by many countries. In Europe, the transition to digital demographic management systems is widespread. Scandinavian nations like Denmark and Sweden have, for many years, conducted population enumeration entirely through administrative registers, providing researchers with high quality, linkable datasets (*United Nations Statistics Division*, 2019; *Juranek – Zoutman*, 2021). Estonia's X-Road platform integrates state information systems and automates data exchange among institutions; by linking register data via unique citizen identifiers, it has enhanced the efficiency and transparency of e-services and even enabled cross-border data sharing with Finland (*Vassil*, 2016; *OECD*, 2021). Guided by recommendations from the United Nations Statistical Commission, countries are urged to modernise demographic data systems through "digital-first" strategies, multi-channel data collection methods (such as mobile applications, telephone surveys, e-forms), and the integration of geographic information systems. Such initiatives will lay the groundwork for more agile, accurate, and comparable demographic statistics in the future (*United Nations Statistical Commission*, 2025). *Weber and State* (2017) define "digital demography", showing how digital trace data can complement surveys/censuses to yield timelier, granular indicators. They also caution about selection bias and privacy, urging triangulation with official statistics (*Weber – State*, 2017).

In recent years, e-demography has emerged as a pivotal approach for analysing population dynamics through the integration of digital data sources and computational methods. The development of e-demography systems has been driven by the expansion of population registers, administrative databases, and other electronic records. These systems provide more timely, detailed, and comprehensive demographic analyses, thereby supporting evidence-based policy and planning. Looking forward, e-demography is expected to advance further through the incorporation of big data, artificial intelligence (AI), and real-time analytics. Such innovations will enhance the capacity of governments and researchers to monitor population trends, evaluate social outcomes, and deliver personalised public services.

POPULATION REGISTERS AS DATA SOURCES FOR E-DEMOGRAPHY SYSTEMS

The primary data backbone of an e-demography system is constituted by government information resources, notably the national population register. According to the *United Nations Statistics Division* (2019), a national population register is an individualised information system that ensures the continuous and coherent recording of selected data for each member of a country's permanent population. Such a register operates on the principle of regularly updating the central database with all demographic events such as births, deaths, marriages, divorces, and changes of residence, thereby enabling real-time access to accurate population counts and structures. In practice, the reliability of register-based systems is closely linked to the strength of civil registration and vital statistics (CRVS), because many core demographic events enter the register through civil-status registration mechanisms.

Primarily, the population register serves to provide reliable information for governmental management and planning purposes. For instance, register data are indispensable for budget allocation, taxation, insurance, social protection, healthcare, education, electoral administration, military conscription, and other routine administrative functions. Furthermore, a comprehensive population register enables rapid and precise identification of citizens (in accordance with the "single window" principle). Experts note that the continuous administrative use of the register also enhances its data quality: daily usage facilitates the prompt detection and correction of errors, thereby establishing a robust statistical foundation (*United Nations Statistics Division*, 2019).

Within the e-demography framework, government registers are integrated not only for administrative objectives but also for statistical analysis. Whereas traditional censuses collected data once every decade, modern population registers allow demographic indicators to be updated with near real-time precision. This capability constitutes a revolutionary advance for agile monitoring of demographic processes and short-term forecasting. It is also important to distinguish the integration of data sources (registers and platforms) from the subsequent statistical evaluation and modelling that may be applied to the integrated data. For example, the Scandinavian experience demonstrates that the coordinated expansion and linkage of individual registration records over many years have provided demographers with unparalleled advantages. *Lyngstad and Skardhamar* (2011) emphasise that demographers in Norway, Sweden, and Denmark benefit most from possessing linked, individual-level demographic databases. Such integrated systems underpin the long-term tracking of demographic trends, cohort analyses, precise measurement of migration flows, and other complex investigations. Similarly, efforts in Belgium and the Netherlands to merge various administrative registers with census data have significantly expanded demographic analysis capabilities (*Lyngstad – Skardhamar*, 2011). These linked registries can be treated as "data sources" that enable particular methodological choices (for example, record linkage, cohort analysis, forecasting, and other computational approaches), which are discussed in the methods-focused sections of this paper.

In the context of Azerbaijan, the integration of government registers is likewise regarded as foundational to e-demography. The Presidential Decree issued in 2004 on the establishment of the "State Population Register of the Republic of Azerbaijan" initiated the creation of a unified population information system. Under the Ministry of Justice's oversight, technological infrastructure was developed to link existing databases across government agencies and to ensure rapid access to necessary information. A central register was established, containing core identification data for all citizens and legal permanent residents – such as full names, patronymics, and dates and places of birth. This central database is updated daily with inputs from sectoral registers (birth and death records, marriage registrations, migration records, etc.). Because population registers are not universal internationally, this national approach can be interpreted as one possible model of e-demography infrastructure rather than a universal prerequisite.

As a result, Azerbaijan's unified population register not only facilitates the "single window" provision of diverse state services but also enables the operational generation of demographic statistics. To date, nearly 17 million archival civil-status records have been digitised and incorporated into the register, and relevant agencies have been equipped with hardware and software to support real-time data transmission. These initiatives constitute the fundamental infrastructure of the national e-demography system. Conceptual studies emphasise that Azerbaijan's e-demography platform must be built upon the unified government population register and that all electronic government registers should integrate into this central framework. Such an approach is strategically vital to ensure that the e-demography system is supplied with reliable and comprehensive data and to facilitate the coordination of demographic information across diverse e-government services. Furthermore, the rapid development of e-government and digital transformations in public administration profoundly influences government–citizen interactions and highlights the importance of providing personalised services (*Yusifov*, 2025). A dedicated section later in the study will provide a brief, structured description of the population register and associated digital platforms as practical enablers for e-demography in Azerbaijan.

REALISATION OF THE ASAN SERVICE BASED ON POPULATION REGISTERS

A practical e-demography framework typically depends on two mutually reinforcing components: (i) authoritative administrative registers that store legally validated demographic attributes and continuously updated life events; and (ii) digital public service platforms that operationalise these registries in citizen–state interaction while generating process information useful for monitoring accessibility and service quality. In this sense, e-demography should be understood not only as the availability of "more data", but as the emergence of an integrated socio-technical setting where data infrastructures and institutional arrangements enable more timely population governance (*Weber – State*, 2017; *Kashyap*, 2021). The Azerbaijani case is relevant for review-based discussion because the country has pursued (a) a unified population register and (b) a standardised one-stop public service model through the Azerbaijan Service and Assessment Network (ASAN), both of which can support e-demography-oriented applications when used with appropriate safeguards.

International statistical guidance defines a population register as an individualised information system that continuously records selected attributes for members of a country's usual resident population and is regularly updated through demographic events, including births, deaths, marriages, divorces, and changes of residence (*United Nations Statistics Division*, 2019). Such systems are not merely technical databases; they function as a governance instrument supporting administrative planning and, under certain conditions, the production of demographic statistics. A register's demographic value emerges from (i) continuity of updates, (ii) standardised definitions, and (iii) the possibility of controlled linkage across domains (*United Nations Statistics Division*, 2019). At the same time, register-based approaches do not eliminate the need for civil registration and vital statistics (CRVS); rather, high-quality CRVS remains essential because it provides the primary institutional channel through which vital events become recorded, validated, and suitable for statistical use (*United Nations Statistics Division*, 2019). This point is also consistent with broader debates in digital demography: even where digital systems expand data availability, the quality of inference depends on coverage, definitions, and governance rules (*Cesare et al.*, 2018; *Wunsch et al.*, 2024).

In Azerbaijan, the creation of a unified population register was initiated through a presidential decision in 2004 that set the direction for establishing a state register of the population and organising inter-agency coordination for population information management (*Presidential Decree of the Republic of Azerbaijan*, 2004). In the Azerbaijani context, digitisation efforts and inter-agency connectivity are also emphasised in national-level discussions on strengthening electronic governance and demographic information systems, where the population register is treated as a central infrastructure for consolidating demographic information and supporting state services (*Alguliyev et al.*, 2019; *Yusifov*, 2021; *Yusifov – Akhundova*, 2022).

A second practical component in Azerbaijan's digital governance landscape is the ASAN Service, established by presidential decree in 2012 as a one-stop model for delivering public services, with objectives centred on improving service coordination, standardisation, transparency, and citizen satisfaction (ASAN Service). While a population register primarily supports the authoritative recording of demographic attributes and events, a service platform such as ASAN represents the operational interface through which citizens access state services and through which the state can observe patterns of service demand and access barriers. This is relevant for e-demography because demographic governance increasingly depends on understanding not only population structure, but also the accessibility and usability of public services across demographic groups especially in contexts where digital transformation is expected to reduce administrative burden and improve inclusion (*Weber – State*, 2017; *Kashyap*, 2021).

Overall, Azerbaijan's population register and the ASAN Service can be presented as practical enablers that illustrate how digital governance infrastructures may support e-demography. In review terms, the key contribution is not an "analysis result", but a structured interpretation of how register-based infrastructures and service platforms can enhance timeliness, coordination, and policy relevance, while requiring explicit methodological discipline and governance safeguards consistent with international guidance and the broader digital demography literature (*United Nations Statistics Division*, 2019; *Kashyap – Zagheni*, 2023).

TECHNOLOGICAL ADVANCES AND ANALYTICAL METHODS IN E-DEMOGRAPHY

E-demography involves not only the digitisation of administrative registries but also the full exploitation of capabilities offered by modern information technologies. Within the e-demography paradigm, the concept of big data is characterised by multiple "V" attributes, including volume, velocity, variety, variability, and visualisation capability (*Nabibayova*, 2021). These attributes demand novel approaches to the collection and processing of demographic data that differ fundamentally from traditional methods.

In recent years, artificial intelligence (AI) and machine learning (ML) algorithms have become indispensable tools for analysing large-scale demographic datasets. For example, clustering techniques (for example, k-means) are employed to uncover latent patterns in high-volume data, thereby enabling segmentation of population subgroups according to shared characteristics (*Yusifov – Akhundova*, 2021). Although methods such as k-means have long been part of established statistical toolkits, their value in e-demography is strengthened when integrated datasets and high-frequency digital sources enable more granular and timely segmentation. Such segmentation allows researchers to examine factors influencing demographic processes within distinct cohorts (for instance, fertility behaviour among different income brackets or migration trends across various regions). Other ML applications include predictive modelling: beyond conventional statistical regression, more sophisticated forecasting models are built using artificial neural networks. In particular, deep learning methods confer unique advantages in extracting actionable insights from unstructured big data.

One of the most compelling applications of deep learning in contemporary demography is population estimation through remote sensing and satellite imagery. Because traditional censuses occur at multi-year intervals, producing an accurate settlement map between enumeration points is challenging. However, by applying deep learning models to high-resolution satellite images and geospatial data, researchers have demonstrated the feasibility of an "alternative census". For instance, *Neal et al*. (2022) applied representation learning algorithms to satellite imagery over Mozambique and generated highly accurate population density maps without reliance on an official census. Their results matched the precision of leading global datasets such as WorldPop and Facebook's High Resolution Settlement Layer (HRSL). These models infer population counts and habitation extents from automatically detected features, such as building footprints, nighttime illumination, and road networks, in satellite imagery. Thus, deep learning algorithms integrate geographic information with demographic statistics to support the creation of continuously updated population databases. Nevertheless, such approaches should be interpreted as complementary tools that require careful validation against official demographic sources and well-defined quality criteria.

The Internet of Things (IoT) technologies also hold significant promise for digital demography. IoT generally denotes the real-time collection of data from various sensors, smart devices, and internet-connected equipment. Within a demographic context, IoT applications span several domains. In urban environments, deployed sensors (for example, smart cameras, traffic counters, mobile-phone towers) enable the monitoring of population mobility and urban flows. Under the Smart City framework, data collected from these devices can quantify the number of individuals in a given area at different times and track their movements, thereby contributing to the real-time monitoring of migration and urbanisation processes. Moreover, wearable devices (fitness trackers, smartwatches) produce large volumes of health and physical activity data that serve as valuable demographic indicators of population well-being. Data from such devices can be aggregated to derive statistics on community-level physical activity, sleep patterns, and cardiovascular metrics, disaggregated by geographical area and age cohort (*Zaman et al.*, 2024). Consequently, the rich datasets harvested from IoT devices enhance demographic analyses, enabling more informed decision-making in public health, transportation planning, and emergency preparedness. As with other digital traces, the interpretability of IoT-derived indicators depends on coverage, representativeness, privacy safeguards, and appropriate aggregation.

DEVELOPING AN E-DEMOGRAPHY SYSTEM: CHALLENGES AND OPPORTUNITIES

One of the principal advantages of the e-demography approach is its ability to ensure the timeliness and responsiveness of demographic data. Unlike the traditional census model, e-demography can support more frequently updated statistics in near real time, depending on the underlying data infrastructure and reporting mechanisms. This capability allows changes in population size and composition (births, deaths, migration, etc.) to be monitored on an annual, monthly, or even daily basis. For example, a system built upon a national population register and supported by strong civil registration and vital statistics (CRVS) can compute indicators such as annual population growth rate, components of natural increase, and migration balance on a monthly schedule (*United Nations Statistics Division*, 2019). Such agility enables governments to implement adaptive demographic policies as issues and trends arise, eliminating reliance on data that are years out of date (*Cesare et al.*, 2018).

Another strength of e-demography lies in its high level of data granularity. Because population registers and integrated digital sources contain individual-level records, analyses can be easily disaggregated by any demographic subgroup (age, sex, region, socio-economic status, etc.). This permits more nuanced examinations of demographic heterogeneity – for instance, calculating birth rates among young families in a specific city or net migration flows in a particular district rather than relying on national averages. By integrating multiple data sources, e-demography systems also facilitate multifactorial analyses of demographic processes. In addition to conventional demographic metrics (births, deaths, marriages, divorces, migration), socio-economic indicators such as educational attainment, income level, health status, and employment can be combined on a single platform, yielding a richer understanding of the causes and consequences of demographic change, which in turn supports evidence-based decision-making (*Alburez-Gutierrez et al.*, 2019). However, since not all countries maintain population registers, e-demography applications must also be framed in relation to classical sources, especially CRVS, censuses, and surveys depending on national statistical capacity (*United Nations Statistics Division*, 2019).

E-demography also eliminates redundancy and enhances efficiency: rather than multiple agencies collecting the same data independently, all government bodies can draw from a single register, reducing the need to request identical information from citizens multiple times. Finally, digital platforms improve the visualisation (interactive maps, dashboards) and dissemination of data, raising public awareness and creating open data resources for researchers (*Alburez-Gutierrez et al.*, 2019). In such cases, transparency and dissemination should be balanced with privacy protection through aggregation and appropriate governance safeguards (*Cesare et al.*, 2018).

Despite these significant benefits, e-demography faces several challenges. The foremost concern is privacy and data security. Population registers and integrated demographic databases contain sensitive personal information on every citizen. Protecting these data against unauthorised access and misuse requires robust cybersecurity measures and comprehensive legal frameworks. Public scepticism may arise when governments collect personal data at scale, making it imperative to communicate the system's benefits and build public trust (*Civelek*, 2023).

Second, data quality and completeness can pose problems. If a birth or death notification is not reported to the register, or if a migrating individual fails to deregister, gaps will emerge in the register data, undermining statistical accuracy. Therefore, the legal and administrative procedures governing data submission must be rigorously enforced, and all relevant agencies must treat data provision as a high priority (*Civelek*, 2023). Routine quality checks and reconciliation with official vital statistics can strengthen consistency and reduce coverage errors (*United Nations Statistics Division*, 2019).

Third, technological integration challenges may arise. Information systems used by different government agencies may lack interoperability, and data formats may vary. Building an e-demography system often requires updating existing infrastructure and implementing interoperability standards, which entails significant time and financial investment. Fourth, there may be a shortage of skilled personnel: specialists such as data scientists and demographic statisticians who can work effectively with big data and AI/ML technologies are essential to ensure that collected data are analysed productively. Without such expertise, even the most advanced technologies may fail to realise their full potential (*Bosco et al.*, 2022).

Methodological challenges are also notable. New digital data sources may not fully represent the entire population; for example, elderly cohorts who use the internet infrequently leave fewer digital traces. This introduces potential bias in big data analyses, and results must be generalised with caution (*Cesare et al.*, 2018). At the same time, algorithmic transparency becomes a concern: complex "black-box" ML models may produce predictions that are difficult to interpret. For demographic policymaking, the explainability of forecasting models is crucial, so that decision-makers can understand the basis of the model's outputs (*Civelek*, 2023). Accordingly, triangulation with official statistics (for example, CRVS, censuses, and surveys) is recommended to improve validity and interpretability (*Weber – State*, 2017). Finally, the lack of international standardisation presents a challenge. As countries develop disparate systems, the comparability of their statistics becomes problematic. Accordingly, methodological guidelines, harmonised definitions, and measurement standards should be established under the auspices of international organisations (*Bosco et al.*, 2022).

Despite these challenges, the benefits of e-demography can outweigh its difficulties when the obstacles are addressed through appropriate legal, technical, and governance measures, and when public support is secured, thereby enabling the full potential of e-demography.

CONCLUSION

This study demonstrates that, within an e-government context, national population registers provide new opportunities for conducting in-depth demographic research. Consequently, the establishment of an e-demography system based on a central population register is essential for the effective management of demographic processes, the analysis of demographic behaviours, and the execution of related research. We reviewed international best practices for developing e-demography systems and analysed research on the operational performance of unified registers. In addition, the Azerbaijani case was discussed as a practical context by outlining the population register and the ASAN Service model as enabling components for e-demography-oriented applications. Note that digital transformation has revitalised the use of both statistical methods and administrative data sources, such as electronic registers, in demographic research. However, the validity of e-demography applications depends on clear separation between the integration/collection of data and subsequent statistical evaluation and modelling, as well as on data quality, representativeness, and governance safeguards. Future research will focus on expanding socio-demographic analyses across diverse domains and applying advanced analytical methods, while ensuring alignment with official demographic sources such as civil registration and vital statistics.

References

- Alburez-Gutierrez, D. – Zagheni, E. – Aref, S. – Gil-Clavel, S. – Grow, A. – Negraia, D. V. 2019. Demography in the digital era: New data sources for population research. In G. Arbia, et al. (Eds.). *Book of short papers*. Pearson. https://doi.org/10.31235/osf.io/24jp7.
- Alguliyev, R. M. – Aliguliyev, R. M. – Yusifov, F. F. – Alekperova, I. Y. 2019. Developing Electronic Demography as an Effective *Tool for Social Research and Monitoring Population Data*. Public Administration Issue, 4, 61–86. (in Russian). https://doi.org/10.17323/1999-5431-2019-0-4-61-86.
- ASAN Service. About ASAN Service (institutional overview and objectives). https://asan.gov.az/en/about.
- Bosco, C. – Grubanov-Boskovic, S. – Iacus, S. – Minora, U. – Sermi, F. – Spyratos, S. 2022. *Data innovation in demography, migration and human mobility* (Preprint No. 2209.05460). arXiv. https://doi.org/10.48550/arXiv.2209.05460.
- Cesare, N. – Lee, H. – McCormick, T. – Spiro, E. – Zagheni, E. 2018. Promises and pitfalls of using digital traces for demographic research. *Demography*, 55(5), 1979–1999. https://doi.org/10.1007/s13524-018-0715-2.
- Civelek, H. Y. 2023. Exploring digital context and its limitations in social sciences: A concentration on demographic research. *Migration and Diversity*, 2(3), 367–381. https://doi.org/10.33182/md.v2i3.3135.
- Grow, A. et al. 2022. Is Facebook's Advertising Data Accurate Enough for Use in Social Science Research? Insights from a Cross-National Online Survey. *Journal of the Royal Statistical Society Series A: Statistics in Society*, 185 (Supplement_2), S343–S363. https://doi.org/10.1111/rssa.12948.
- Juranek, S. – Zoutman, F. T. 2021. The effect of non-pharmaceutical interventions on the demand for health care and on mortality: Evidence from COVID-19 in Scandinavia. *Journal of Population Economics*, 34(4), pp. 1299–1320. https://doi.org/10.1007/s00148-021-00868-9.
- Kashyap, R. 2021. Has demography witnessed a data revolution? Promises and pitfalls of a changing data ecosystem. *Population Studies, 75*(sup1), 47–75. https://doi.org/10.1080/00324728.2021.1969031.
- Kashyap, R. – Zagheni, E. 2023. Leveraging digital and computational demography for policy insights. In E. Bertoni et al. (Eds.). *Handbook of Computational Social Science for Policy*. Cham: Springer. Pp. 327–344. https://doi.org/10.1007/978-3-031-16624-2_17.
- Lyngstad, T. H. – Skardhamar, T. 2011. Nordic register data and their untapped potential for criminological knowledge. *Crime and Justice*, 40(1), 613–645.
- Nabibayova, G. 2021. Application of OLAP technology in the environment of electronic demography. https://doi.org/10.17587/it.27.542-549.
- Neal, I. – Seth, S. – Watmough, G. – Diallo, M. S. 2022. Census-independent population estimation using representation learning. *Scientific Reports, 12*(1), 5185. https://doi.org/10.1038/s41598-022-08935-1.
- OECD. 2021. X-Road Trust Federation for Cross-border Data Exchange. *OECD Digital Government Studies*. Paris: OECD Publishing.
- Poulain, M. – Herm, A. 2013. Central Populatıon Regısters as a Source of Demographic Statıstıcs in Europe. *Population*, vol. 68, No 2, 183–212. https://doi. org/10.3917/popu.1302.0215.
- Presidential Decree of the Republic of Azerbaijan. 2004. *Decree on creation of the State Register of Citizens of the Azerbaijan Republic*. https://e-qanun.az/framework/4808 (in Azerbaijani).
- United Nations Statistical Commission. 2025. *Towards an overarching conceptual framework for social and demographic statistics*. Fifty-sixth session, 4-7 March 2025.
- United Nations Statistics Division. 2019. Population registers and their uses. In *Principles and Recommendations for Population and Housing Censuses*, Rev. 3, paras. 454-482. New York: United Nations.
- Vassil, K. 2016. *Estonian e-government ecosystem: Foundation, applications, outcomes*. World Bank. https://doi.org/10.1596/27702.

- Weber, I. – State, B. 2017. Digital demography. *In Proceedings of the 26th International World Wide Web Conference (WWW 2017)*, Perth, Australia: ACM. pp. 935–939. https://doi.org/10.1145/3041021.3051104.
- Wilde, J. – Chen, W. – Lohmann, S. 2020. COVID-19 and the future of US fertility: What can we learn from Google? *IZA Discussion Paper*, No. 13776. Bonn: Institute of Labor Economics (IZA).
- Wunsch, G. – Gourbin, C. – Russo, F. 2024. Big data, demography, and causality. Open *Journal of Social Sciences*, 12(1), 181–206. https://doi.org/10.4236/jss.2024.121012.
- Yusifov, F. 2021. Using Public Registers for Development of Electronic Demography System: The Case of Azerbaijan. *International Journal of Computing and Digital Systems*, 1, 1251–1261. https://doi.org/10.12785/ijcds/1001112.
- Yusifov, F. 2025. A citizen-centric approach for digital government transformation: The case of Azerbaijan. *Information Development*. Online first. https://doi.org/10.1177/02666669251370361.
- Yusifov, F. F. – Akhundova, N. E. 2021. Analysis of demographic indicators based on the data of the e-demography system. *Problems of Information Society*, 12(2), 70–82. https://doi.org/10.25045/jpis.v12.i2.05 (in Azerbaijani).
- Yusifov, F. F. – Akhundova, N. E. 2022. Analysis of demographic characteristics based on e-demography data. *Demohrafiia ta sotsialna ekonomika - Demography and Social Economy*, 1(47), 38–54. https://doi.org/10.15407/dse2022.01.038.
- Zagheni, E. – Weber, I. – Gummadi, K. P. – State, B. 2019. Facebook advertising data for demography: Joint uses of survey, census, and social media data for population research. *Population and Development Review*, 45(1), 130–157.
- Zaman, M. – Puryear, N. – Abdelwahed, S. – Zohrabi, N. 2024. A review of IoT based smart city development and management. *Smart Cities*, 7(3), 1462–1501. https://doi.org/10.3390/smartcities7030061.

FARHAD YUSIFOV

Received his PhD degree in Information Technology in 2010 from the Institute of Information Technology in Baku, Azerbaijan. His primary research interests include digital transformation, digital government, e-governance, and e-demography, with a particular focus on data analytics and multi-criteria decision-making. He currently serves as the head of the Digital Transformation and E-Government Department at the Institute of Information Technology. He is the author of over 80 scholarly publications.

ELVIN GURBANOV

Is a scientific researcher at the Institute of Information Technology in Baku, Azerbaijan. He received his BSc degree with honours from the Azerbaijan State University of Economics (UNEC) and his MSc degree from the Azerbaijan University of Architecture and Construction (AUAC). His primary research interests include e-demography, digital transformation, digital government, and e-governance. He currently serves as a programme coordinator in Computer Science and as a lecturer at Western Caspian University and AUAC.

REGIONÁLNÍ DIFERENCIACE REPRODUKČNÍHO CHOVÁNÍ V ČESKU V LETECH 2015, 2019 A 2023: DESKRIPTIVNÍ A SHLUKOVÁ ANALÝZA

Kristýna Dvořáková[1)] – Filip Hon[2)]

REGIONAL DIFFERENTIATION OF REPRODUCTIVE BEHAVIOUR IN CZECHIA IN 2015, 2019, AND 2023: DESCRIPTIVE AND CLUSTER ANALYSIS

Abstract

This article examines regional differences in reproductive behaviour across Czech districts in 2015, 2019, and 2023. Its aim is to describe spatial disparities in selected indicators, assess their temporal stability, and identify groups of districts with similar reproductive patterns. The analysis draws on data from the Demographic Yearbook of Districts published by the Czech Statistical Office and uses descriptive statistics, correlation analysis, and hierarchical cluster analysis based on components derived through principal component analysis. The results confirm there is substantial spatial variability in the observed indicators as well as stable relationships between them. The cluster analysis identified five groups of districts, with the Northwestern Region forming the most internally homogeneous cluster, characterised by a low maternal age at first birth, a high share of nonmarital births, elevated levels of induced abortions for non-medical reasons, and a higher proportion of third- and higher-order births. The remaining clusters represent distinct types of reproductive behaviour, collectively demonstrating the considerable diversity of reproductive patterns across Czechia.

Keywords: Reproductive behaviour, Regionalisation, Cluster analysis
Demografie, 2026, **68(1): 52–71**
DOI: https://doi.org/10.54694/dem.0379

ÚVOD

Reprodukční chování obyvatel Česka v posledních letech vykazuje výrazné výkyvy. Například při pohledu na celkovou úroveň plodnosti v období celosvětové pandemie onemocnění Covid-19 dosáhla úhrnná plodnost v roce 2021 hodnoty 1,83, již v roce 2024 však úhrnná plodnost poklesla na 1,37 (*ČSÚ*, 2025).

1) Vysoká škola ekonomická v Praze, Fakulta informatiky a statistiky. Kontakt: dvok17@vse.cz.
2) Vysoká škola ekonomická v Praze, Fakulta informatiky a statistiky. Kontakt: filip.hon@vse.cz.

Tento výrazný propad se stal předmětem nejen odborných diskusí, ale také mediálního zájmu, a téma demografického vývoje tak rezonuje stále výrazněji i v širší veřejnosti.

Pokles úrovně plodnosti je často vysvětlován kombinací faktorů, mezi něž patří zejména ekonomická nejistota, vliv mezinárodních konfliktů či zhoršující se environmentální situace (*Paloncyová a kol*., 2021; *Waldaufová – Šťastná*, 2025). Tyto faktory se promítají do rozhodovacího procesu jednotlivců a domácností o zakládání rodiny a počtu dětí.

Reprodukční chování v tomto článku chápeme jako soubor demografických procesů a rozhodování souvisejících s realizací rodičovství, který zahrnuje nejen intenzitu plodnosti, ale také načasování vstupu do mateřství, strukturu porodů podle pořadí, rodinný kontext početí a narození dítěte a způsoby reprodukční regulace, včetně výskytu umělých přerušení těhotenství (zkratka UPT). Úhrnná plodnost přitom představuje pouze jeden syntetický indikátor, který nepostihuje zásadní změny probíhající v české společnosti. Současné reprodukční chování je výrazně utvářeno i dalšími procesy.

Časování plodnosti v Česku se dlouhodobě posouvá do vyššího věku, mezi zamýšleným a skutečným věkem při narození prvního dítěte se vytváří výrazná mezera. Odklad prvního porodu významně snižuje prostor pro další děti. Tento rozdíl je způsoben nejen individuálními rozhodnutími, ale také strukturálními faktory, jako je prodlužující se vzdělávání a nestabilita partnerských vztahů. Většina české populace stále preferuje dvoudětný model rodiny a za ideální věk pro vstup do rodičovství považuje dobu před třicátými narozeninami. Reálné reprodukční plány však zůstávají skromnější. Důvodem jsou přetrvávající ekonomické nejistoty, vysoké ceny bydlení a nedostatečná institucionální podpora rodin. Mladší generace zároveň čelí řadě překážek, mezi něž patří obtížnost najít stabilního partnera, časová náročnost péče o děti a také širší nabídka možností osobního rozvoje, které často soutěží s rodičovstvím jako životní prioritou (*Slabá a kol*., 2024; *Waldaufová – Šťastná*, 2025).

Dále je utvářeno například mimomanželským rodičovstvím, jehož nárůst je jedním z klíčových rysů post-socialistické transformace. Odborné studie ukazují, že podíl dětí narozených mimo manželství v Česku vzrostl v posledních desetiletích několikanásobně, především v souvislosti s posilující individualizací a změnami hodnotových orientací (*Stípková*, 2015).

Podobně významnou součástí reprodukčního chování je i potratovost, zejména intenzita záměrně vyvolaných přerušení těhotenství, která doplňuje informaci o reprodukčních preferencích, normách a rodinných strategiích. Přehledové analýzy ukazují, že česká potratovost se v posledních letech stabilizovala na nízké úrovni (*Kocourková a kol*., 2023).

Důležitým faktorem při hodnocení reprodukčního chování je také to, jak časté jsou v populaci rodiny se třemi a více dětmi a jakým způsobem se promítají do reprodukčních drah jednotlivých generací. Dostupné analýzy ukazují, že míry plodnosti třetího pořadí se v posledních dekádách sice mírně snížily, avšak zůstávají stabilnější než plodnost nižších pořadí a neprošly šokovým propadem v 90. letech. Navíc se posouvají do vyššího věku žen klidně až k 35 letům, zatímco v roce 1990 vrcholila plodnost třetího pořadí kolem 28 let (*Hon*, 2020).

Všechny tyto souvislosti ukazují, že pro pochopení regionální diferenciace reprodukčního chování je nezbytné zahrnout nejen úhrnnou plodnost, ale i ukazatele mimomanželské plodnosti, plodnosti vyšších pořadí, časování vstupu do rodičovství a také intenzitu UPT. Jedná se o ukazatele, které jsou na okresní úrovni dostupné a jejich kombinace umožňuje zachytit komplexní a v posledních letech rychle se měnící podobu reprodukčních procesů v českých okresech podstatně lépe, než pohled pouze na některý konkrétní ukazatel.

Cílem tohoto článku je rozšířit známé skutečnosti o nedávném vývoji reprodukčního chování v Česku o detailní regionální analýzu na úrovni okresů, a to s využitím moderních vícerozměrných statistických metod. Podle našich znalostí podobně podrobná aktuální regionální studie dosud v odborné literatuře není dostupná.

Konkrétněji lze definovat tři hlavní cíle tohoto článku. Prvním cílem je detailně popsat regionální diferenciaci reprodukčního chování v okresech Česka na základě vybraných demografických ukazatelů. Druhým cílem je analyzovat, jak se tyto regionální vzorce změnily mezi lety 2015, 2019 a 2023, a posoudit stabilitu či dynamiku sledovaných charakteristik v čase. Třetím cílem je identifikovat skupiny okresů s obdobnými reprodukčními vzorci pomocí shlukové analýzy a charakterizovat jejich specifické rysy. Studie tak přispívá k hlubšímu pochopení prostorového rozložení reprodukčních vzorců v Česku a vytváří rámec pro následnou diskusi o jejich potenciálních socioekonomických a kulturních souvislostech.

Výzkumné otázky jsou na základě uvedených cílů definovány následovně. Článek zkoumá, jaké regionální rozdíly v reprodukčním chování okresů Česka lze pozorovat v letech 2015, 2019 a 2023 na základě vybraných ukazatelů. Text dále porovnává, jak se tyto rozdíly měnily ve sledovaném čase a do jaké míry vykazují sledované ukazatele stabilitu či dynamiku mezi uvedenými roky. Vícerozměrná analýza následně řeší problematiku seskupení okresů pomocí shlukové analýzy, a zároveň na základě jakých konkrétních zákonitostí z hlediska reprodukčního chování jsou okresy členěny do právě takových shluků.

Příspěvek je rozčleněn do několika částí, nejprve je zde uvedena rešerše seznamující zejména se staršími regionálními analýzami, jejich zjištěními a obecným přehledem využití vícerozměrné metody shlukové analýzy k regionalizaci v demografii. Druhá část článku je věnována metodologii ohledně dat i následně popisu konkrétní realizace shlukové analýzy, která samozřejmě právě zdrojová data využívá při výpočtu. Nejrozsáhlejší jsou následně analytické části, kde první zahrnuje deskriptivní statistiku využitých indikátorů reprodukčního chování a druhá se již věnuje výsledkům shlukové analýzy.

REŠERŠE

Deskriptivní i shluková analýza reprodukčních charakteristik okresů Česka není v odborné literatuře novým postupem, naopak představuje etablovanou součást regionální demografie. Shluková analýza dokonce umožňuje identifikovat územní jednotky s obdobnými demografickými rysy a vytvářet tak skupiny s vnitřní homogenitou a vnější odlišností. Z metodologického hlediska lze tento postup chápat jako specifický typ regionalizace, tedy procesu, při němž jsou územní celky členěny do regionů na základě sdílených charakteristik či vzorců chování (*Fialová*, 2018). Regionalizace v tomto pojetí nepředstavuje pouhé geografické vymezení, ale analytický nástroj zachycující prostorovou strukturu demografických jevů.

Regionalizace pomocí zvolených metod v tomto článku je pak běžná v demografii obecně, nejen v případě okresů. Shluková analýza je využívána nejen v kontextu Česka, ale i na mezinárodní úrovni, například příspěvek s názvem *Time Series Clustering Of Fertility in European Union* se zabývá shlukováním členských zemí Evropské unie na základě vývoje úhrnné plodnosti mezi lety 1965–2021 (*Bakuncová – Löster*, 2023).

K tvorbě shlukové analýzy jsou kromě indikátorů plodnosti, průměrného věku matek při porodu, případně i potratovosti, využívány také jiné demografické charakteristiky. Například Ondřej Šimpach a Marie Šimpachová Pechrová v příspěvku *Cluster Analysis Of EU's Regions According to Demographics Criteria* pomocí dalších ukazatelů, mezi které patří mimo jiné střední délky života, počty zemřelých či populace 85letých a starších, sestavili shlukovou analýzu napříč administrativními jednotkami NUTS2 členských států EU na základě dat z let 2004 a 2012. To dokládá, že vícerozměrné metody jsou standardním nástrojem studia regionální demografie a dále také, že dlouhodobé regionální disparity jsou doložené i v jiných demografických procesech (*Šimpach – Šimpachová Pechrová*, 2016).

Na okresní úrovni existuje řada významných studií, které se v minulosti zabývaly prostorovou diferenciací plodnosti. Následující odstavce shrnují klíčová zjištění, jež k tomuto tématu dosavadní odborná literatura přináší.

Vývoj regionální diferenciace plodnosti v devadesátých letech zkoumá například Bartoňová (1999) ve článku *Vývoj regionální diferenciace věkové struktury se zřetelem k územním rozdílům ve vývoji reprodukce v České republice* rozlišuje čtyři typy okresů v Česku podle vývoje úhrnné plodnosti na přelomu 80. a 90. let. První typ zahrnuje okresy s dlouhodobě podprůměrnou plodností, kde po roce 1992 dochází k výraznému poklesu; typické jsou pro velká města. Druhý typ odpovídá celostátnímu průměru, s pozvolným poklesem plodnosti do začátku 90. let a jejím rychlým snižováním po roce 1993. Třetí typ vykazuje nadprůměrnou úhrnnou plodnost, přestože jeho vývoj jinak odpovídá typu druhému; zahrnuje zázemí měst, průmyslové pohraničí i zemědělské okresy. Čtvrtý typ tvoří převážně zemědělské okresy s vysokou plodností přetrvávající ještě v roce 1991, která se výrazněji snižuje až po roce

1993. Studie dále upozorňuje na podstatně častější mimomanželskou plodnost v severovýchodní části republiky již od roku 1990 (*Bartoňová*, 1999). Obecně lze nicméně konstatovat, že změny reprodukčního chování v devadesátých letech nevedly k výraznějšímu rozevírání regionálních rozdílů v úrovni plodnosti. Spíše než prohlubování diferenciace docházelo k postupné proměně a přeskupování prostorového rozložení plodnosti, tedy k určitému vyrovnávání rozdílů mezi regiony (*Bartoňová*, 2001). Tento vývoj potvrzují také analýzy zaměřené na generační perspektivu, které poukazují na útlum regionálních kontrastů i při sledování rozdílů mezi kohortami (*Rychtaříková*, 2007).

Šídlo (2008) ve svém článku *Faktory ovlivňující regionální diferenciaci plodnosti v Česku na počátku 21. století* zdůrazňuje několik klíčových prostorových vzorců. Článek poskytuje důležitý kontext pro porozumění regionálním rozdílům plodnosti v Česku před nástupem výraznější rekuperace plodnosti a diverzifikace reprodukčního chování, kterou autor popisují v pozdějších pracích. Podle autora se tehdejší výsledky analýz shodují s „všeobecně vnímaným pohledem" na regionální uspořádání plodnosti v Česku. Studie konstatuje postavení městských okresů, které se v rámci zkoumaného období jasně vymezují jako specifická skupina z hlediska reprodukčního chování. Současně se jako problematický region ukazuje severozápadní část Česka. Dalším výrazným zjištěním studie je zachování tradičních kulturních norem v určitých částech země, tyto tradiční vzorce jsou typické zejména pro Vysočinu a střední Moravu, které se tak odlišují od jiných regionů svým konzervativnějším charakterem reprodukčního chování (*Šídlo*, 2008).

Podrobněji je následně vhodné podívat se na některé studie vydané blíže současnosti. Kurkin ve své disertační práci *Faktory plodnosti a regionální diferenciace plodnosti v České republice po roce 1991* pomocí shlukové analýzy potvrzuje, že vývoj úhrnné plodnosti v českých okresech není homogenní, ale vytváří několik relativně stabilních regionálních vzorců. Výrazně se vyčleňuje severozápadní část Česka, která tvoří specifický shluk se zvýšenou intenzitou plodnosti. Shluk s okresy na jižní Moravě prošel přesunem z nadprůměrné plodnosti k podprůměrné. Další výrazný vzorec představují regiony v zázemí Prahy, které přešly z průměrných hodnot k nadprůměrné plodnosti, a rovněž některé okresy jihozápadních a západních Čech, jež se z původně podprůměrných hodnot postupně konvergovaly směrem k průměru. Tyto výsledky ukazují, že český prostor se v období 1991–2014 diferencuje do několika jasně profilovaných trajektorií (*Kurkin*, 2015).

Ve studii *Regionálne rozdiely v charaktere plodnosti v Česku a na Slovensku* autoři analyzují prostorovou diferenciaci rekuperace plodnosti a ukazují, že právě časování mateřských a rodičovských startů je klíčovým faktorem rostoucí variability plodnosti mezi regiony. Zatímco samotná intenzita plodnosti v Česku vykazuje trend prostorového sbližování, časování porodů se naopak stále více rozchází. Tento proces vede k výrazné diverzifikaci reprodukčních modelů a ke vzniku pluralitních vzorců reprodukčního chování. Autoři tak identifikují silnou prostorovou divergenci v načasování rodičovství, která pravděpodobně bude ovlivňovat podobu plodnosti i v budoucích letech. Regiony se tak stále jasněji odlišují podle toho, jak rychle či pomalu přistoupily k posunu rodičovství do vyššího věku (*Šprocha – Šídlo*, 2016a).

Uvedené skutečnosti dále rozšiřuje článek s názvem *Spatial Differentiation and Fertility Postponement Transformation in Czechia* člení okresy Česka do čtyř shluků. Z výsledků vyplynulo, že nízká míra rekuperace po výrazném odkladu plodnosti je hlavní příčinou toho, že většina moravských okresů zaostávala za celostátním průměrem. Transformace plodnosti v Česku zásadně změnila prostorové vzorce reprodukčního chování, i když některé dlouhodobé charakteristiky – zejména ve velkých městských okresech, ekonomických centrech a západních a severozápadních pohraničních regionech – zůstávají stabilní. V západních a severozápadních regionech je celková plodnost vyšší, přičemž výrazněji se to projevuje u mladších žen. Článek ukazuje, že mnoho moravských okresů patří k těm, které se nejhůře přizpůsobily novým reprodukčním podmínkám. Původní jednotný model časného rození dětí však již neplatí a byl převážně opuštěn. Tempo a věková struktura plodnosti, úzce spojené s odkladem a následnou rekuperací, se nyní staly klíčovými prostorovými faktory, které rozlišují reprodukční chování mezi českými okresy. Dochází také k určité konvergenci okresů v kontextu rekuperace (*Šprocha – Šídlo*, 2016b).

V článku *Odkládání mateřství a regionální diferenciace plodnosti v Česku a na Slovensku* je dokumentována výrazná proměna prostorového rozložení plodnosti v Česku. Zatímco začátkem 90. let se vyšší plodnost soustřeďovala zejména v moravských okresech, nejnižší hodnoty vykazovala Praha, její zázemí a další hospodářská centra, následně dochází k inverzi tohoto vzorce. Regiony s původně nejnižší plodností se vyznačovaly nejvýraznějším oživením plodnosti ve vyšším věku a patřily k okresům s nejvyšší celkovou plodností. Tento obrat autoři interpretují jako důsledek rychlého přijetí odloženého rodičovství v ekonomicky silných a urbanizovaných oblastech.

Naopak mnohé moravské regiony, dříve charakteristické relativně vysokou plodností, vykazovaly perzistenci velmi nízké plodnosti, především kvůli absenci významnějšího zvýšení reprodukce ve vyšším reprodukčním věku. Mezi oblasti s nízkou mírou oživení plodnosti patří zejména socioekonomicky problémové regiony. Jde především o příhraniční okresy na jihu a jihovýchodě při hranici se Slovenskem a Rakouskem a dále o příhraniční území na západě a severovýchodě Česka. Autoři upozorňují, že tato území dlouhodobě vykazují slabší socioekonomické ukazatele, což se projevuje také v opožděné či omezené adaptaci na nové reprodukční vzorce (*Šídlo – Šprocha*, 2018).

METODOLOGIE – ZDROJE DAT

Tento článek vychází z Demografické ročenky okresů, která je pravidelně publikována Českým statistickým úřadem (*ČSÚ*, 2024). Analýza se zaměřuje na demografické údaje za roky 2015, 2019 a 2023.

S ohledem na specifické postavení hlavního města Prahy, které funguje zároveň jako kraj i okres, nebyla tato jednotka do analýzy zahrnuta. Zahrnutí Prahy by vzhledem k její demografické výjimečnosti mohlo zkreslit výsledky jako statistické odlehlé pozorování. Tato skutečnost byla prokázána výpočtem Mahalanobisovy vzdálenosti, její hodnoty a podrobnější zdůvodnění jsou následně zobrazeny dále v první analytické části článku.

Analýza tak zahrnuje 76 okresů Česka.

Pro analýzu byly pro všechny okresy Česka (vyjma Prahy) vybrány následující proměnné:

- úhrnná plodnost,
- podíl živě narozených dětí ve třetím a vyšším pořadí,
- podíl živě narozených dětí mimo manželství,
- index uměle přerušených těhotenství (nevynucených zdravotními důvody) a počtu živě narozených (dále jen index UPT),
- průměrný věk matky při narození prvního živého dítěte.

Výběr těchto proměnných byl veden dvěma hlavními kritérii. Důležitá byla jejich dostupnost, tedy skutečnost, že všechny tyto ukazatele jsou ve sledovaných letech k dispozici na úrovni okresů.

Dále byly zařazeny ukazatele rozšiřující kontext reprodukčních návyků obyvatel. Například kromě celkové úrovně aktuální plodnosti charakterizované syntetickým ukazatelem úhrnné plodnosti je tak zahrnuto procento dětí narozených mimo manželství a podobně.

Význam jednotlivých proměnných pro analýzu, jejímž cílem je charakterizovat reprodukční chování, a tedy důvod zařazení těchto proměnných, lze shrnout následovně:

Úhrnná plodnost – charakterizuje tendenci populace mít potomky a celkovou úroveň intenzity plodnosti.

Podíl dětí živě narozených ve třetím a vyšším pořadí – ukazuje, jak často dochází k překročení hranice reprodukce (více než dvě děti na ženu), respektive plodnosti vyšších pořadí.

Podíl dětí živě narozených mimo manželství – odráží proměny rodinných modelů a postoj společnosti k manželství jako předpokladu rodičovství.

Index UPT – tento ukazatel byl vypočten jako podíl, kde v čitateli je počet UPT (s odečtením UPT ze zdravotních důvodů) a ve jmenovateli počet živě narozených dětí. Reprezentuje přerušená těhotenství s vyloučením umělých přerušení ze zdravotních důvodů. Pro lepší interpretaci byl vynásoben stem.

Průměrný věk matky při narození prvního živého dítěte – vypovídá o načasování rodičovství, které souvisí mimo jiné se vzděláním a profesními ambicemi žen.

Struktura použitých ukazatelů zahrnuje jeden syntetický indikátor celkové úrovně reprodukčního chování v podobě úhrnné plodnosti, ukazatel zachycující časování mateřství (průměrný věk prvorodičky) a tři podílové charakteristiky, které doplňují informace o reprodukčních vzorcích v jednotlivých regionech (podíl živě narozených mimo manželství, podíl živě narozených ve třetím a vyšším pořadí a index UPT).

Tyto ukazatele společně poskytují komplexní pohled na intenzitu, načasování i strukturu reprodukčního chování v rámci okresů.

METODOLOGIE – KVANTITATIVNÍ METODY

V úvodu bylo nutné učinit rozhodnutí o (ne)vyřazení některých pozorování (okresů) z analýzy.

Mahalanobisova vzdálenost byla použita k identifikaci okresů, které se ve vícerozměrném prostoru sledovaných demografických ukazatelů odlišují od zbytku souboru. Pro každý okres byla vypočtena vzdálenost:

$$D^2 = (x - \mu)^{\mathrm{T}} S^{-1}(x - \mu), \tag{1}$$

kde x je vektor hodnot daného okresu, μ je vektor průměrů všech okresů a S^{-1} je inverzní kovarianční matice, což zohledňuje nejen rozdílné měřítko jednotlivých ukazatelů, ale také jejich vzájemné korelace, a proto umožňuje posoudit odlehlost okresu v kontextu celého vícerozměrného prostoru (*Pecáková*, 2011). Statistika D^2 je porovnána s chí-kvadrát rozdělením o p stupních volnosti, kde p představuje počet použitých ukazatelů. To umožňuje stanovit hranici, nad kterou lze okres považovat za statisticky signifikantně odlehlý. Kritická hodnota je dána:

$$D^2 > \chi^2_{p,1-\alpha}. \tag{2}$$

V tomto článku byla použita hladina významnosti . Okresy, jejichž Mahalanobisova vzdálenost překročila odpovídající kritickou hodnotu chíkvadrát rozdělení, byly označeny za vícerozměrné odlehlé pozorování (outliers). Tyto jednotky se neodlišují pouze v jedné proměnné, ale v kombinaci všech sledovaných ukazatelů a nelze takto identifikované odlehlé pozorování využít s ohledem na výpočetní algoritmus vícerozměrné metody.

Z hlediska popisné statistiky článek pozoruje kvantily, konkrétně kvartily, jednotlivých proměnných během sledovaných let. Pro účely grafického zobrazení výsledků byly vytvořeny, jak je v podobných případech obvyklé, krabičkové grafy (boxploty), které kromě kvartilů datových souborů zobrazují minimální a maximální neodlehlé hodnoty (whiskers), či jednorozměrná odlehlá pozorování. Odlehlá pozorování jsou taková pozorování, pro která platí jedna z následujících podmínek:

hodnota < první kvartil – (1,5 × mezikvartilové rozpětí),

nebo:

hodnota < první kvartil + (1,5 × mezikvartilové rozpětí),

kde mezikvartilové rozpětí je rovno rozdílu třetího a prvního kvartilu datového souboru (*Řezanková*, 2024).

Dalším zkoumaným rozměrem byla jednorozměrná variabilita vstupních proměnných a její vývoj napříč lety. K tomu byly pro analýzu zkonstruovány variační koeficienty, jejichž výhodou je jejich bezrozměrnost, a tedy možnost srovnání nejenom stejných proměnných v různých časech, ale také různých proměnných (s různými škálami). Jedná se o poměr směrodatné odchylky a průměru vynásobený 100 (díky vynásobení jsou hodnoty uváděny v procentech) (*Hindls a kol.*, 2007).

Deskriptivní analytická část je následně doplněna o přehled korelací zvolených proměnných a jejich vývoj v čase. K výpočtům monotónních závislostí byla použita metoda Spearmanova korelačního koeficientu, jehož výpočet vychází z transformace proměnných na ordinální typ a následného porovnání pořadí dvou sledovaných proměnných pro dané pozorování. Výsledky jsou zobrazeny ve vizualizované podobě korelační matice.

Shluková analýza je jednou z vícerozměrných metod, která se využívá k identifikování struktury v datech, konkrétně k seskupení objektů do skupin, shluků, na základě podobnosti proměnných. Z toho vyplývá, že objekty vyskytující se ve stejném shluku jsou homogenní, naopak objekty z různých shluků jsou heterogenní.

U shlukování není zařazení do skupin předem známé, metoda vytváří nové shluky. Přístupy ke shlukové analýze se z hlediska základního členění dělí na hierarchické a nehierarchické přístupy, u hierarchických přístupů autoři dopředu neurčují počet shluků, u nehierarchických ano.

Hierarchické přístupy se dále dělí na aglomerativní, kdy každý objekt nejprve tvoří samostatný shluk, a později dochází na základě podobností k seskupování objektů ve vícečlenné shluky (v některých případech shluk může stále zůstat jednočlenný), a divizivní, přičemž tyto přístupy uvažují, že všechny objekty zpočátku tvoří jeden shluk, a později dochází k jejich rozdělování na méně početné shluky (*Řezanková*, 2024: 78).

V následující shlukové analýze bylo použito aglomerativní hierarchické shlukování, konkrétně Wardova metoda. Vzhledem k povaze metody byla jako míra podobnosti zvolena čtvercová euklidovská vzdálenost. Wardova metoda patří mezi nejčastěji využívané postupy v rámci hierarchických metod, protože, zjednodušeně řečeno, minimalizuje vnitroshlukovou variabilitu, je poměrně robustní a vytváří kompaktní shluky (*Hebák a kol.*, 2015).

Vzhledem k silným závislostem mezi některými z dvojic proměnných nebyla shluková analýza konstruována z těchto dat, nýbrž z komponent. Komponenty byly vytvořeny pomocí metody hlavních komponent, a to na základě jejich hlavní vlastnosti, tedy vzájemné kolmosti, resp. nezávislosti. Tím se liší od jiných přístupů, jako je např. výpočet z-skórů, kdy takový přístup sice znormalizuje vstupní data, avšak proměnné mohou i nadále být korelované. Komponenty vycházejí z původních dat výše zmiňovaných proměnných za 3 roky – 2015, 2019 a 2023 (dohromady je analýza tvořena 15 proměnnými). Vzhledem k rozdílným škálám i rozptylům původních proměnných, transformace vstupních dat metodou hlavních komponent vycházela z korelační matice. Byť tato metoda bývá často využívána jako redukční algoritmus, pro současnou úlohu nebylo potřeba snižovat počet vstupujících proměnných a redukovat tak vstupní informaci. Motivací využití metody hlavních komponent je skutečnost, že tvorbou nových nekorelovaných latentních proměnných pomáhá metoda s fungováním shlukových algoritmů (*Jolliffe*, 2002).

Výpočty a vizualizace analýzy byly provedeny pomocí softwarů MS Excel *(Microsoft Corporation*, 2024), SPSS (SPSS Inc., 2009) a programovacího jazyku R (*R Core Team*, 2024). Konkrétně byly využity následující balíčky: *corrplot* (*Wei – Simko*, 2024), *dplyr* (*Wickham a kol.*, 2023), *ggplot2* (*Wickham et al.*, 2025), *RCzechia* (*Lacko*, 2024), *readxl* (*Wickham – Bryan*, 2025), *sf* (*Pebesma*, 2018), *tidyverse* (*Wickham a kol.*, 2024).

DESKRIPTIVNÍ ANALÝZA VSTUPNÍCH DAT

Cílem kapitoly je poskytnout detailní přehled o základních rysech regionální diferenciace reprodukčního chování a identifikovat hlavní okruhy, ve kterých jsou rozdíly mezi administrativními jednotkami nejvýraznější. Kapitola se zaměřuje na to, jak se jednotlivé sledované proměnné v okresech liší a jak se jejich prostorové rozložení proměňovalo mezi lety 2015, 2019 a 2023.

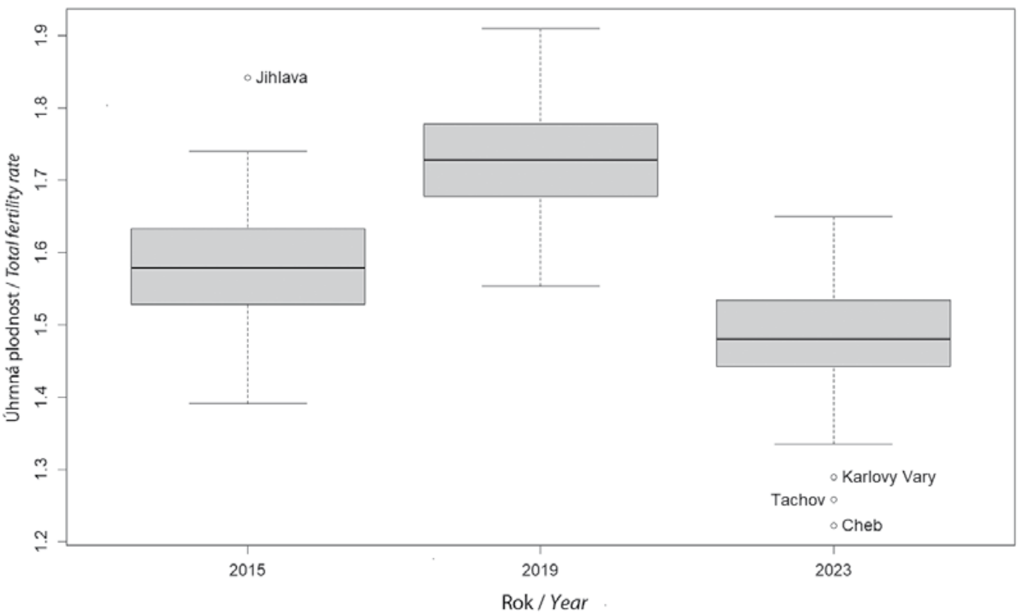
Tato část tak vytváří analytický základ pro následné komplexnější porovnání okresů a pro vysvětlení podobnosti jejich reprodukčních profilů pomocí vícerozměrné statistiky. Deskriptivní statistika a přehled monotónních vztahů mezi ukazateli zde slouží jako nástroj k pochopení toho, jaké charakteristiky českých

Tab. 1: Vícerozměrná odlehlá pozorování v letech 2015, 2019 a 2023 pomocí výpočtu Mahalanobisovy vzdálenosti / Multivariate outlier detection in the years 2015, 2019, and 2023 calculated using Mahalanobis distance

2015		2019		2023	
Okres / *District*	**Mahalanobisova vzdálenost *Mahalanobis distance***	**Okres / *District***	**Mahalanobisova vzdálenost *Mahalanobis distance***	**Okres / *District***	**Mahalanobisova vzdálenost *Mahalanobis distance***
Hlavní město Praha *Capital City of Prague*	28,7	Hlavní město Praha *Capital City of Prague*	25,3	Hlavní město Praha *Capital City of Prague*	27,5
Praha-západ *Prague-West*	15,5	Jeseník	21,2	Jeseník	20,7
Jihlava	13,3	Ústí nad Labem	14,1	Cheb	14,4
		Žďár nad Sázavou	12,9	Klatovy	13,3
				Praha-západ *Prague-West*	13,1

Zdroj: Vlastní zpracování na základě dat ČSÚ.
Source: *Authors' calculation based on CZSO data.*

Graf 1: Regionální diferenciace úhrnné plodnosti v letech 2015, 2019 a 2023 (soubor okresů bez Prahy) – krabičkový graf / Regional differentiation of the total fertility rate in 2015, 2019, and 2023 (a set of districts without Prague) – box plot



Zdroj: Vlastní zpracování na základě dat ČSÚ.
Source: *Authors' calculation based on CZSO data.*

okresů se v čase sbližují, které se naopak rozcházejí a v jakých demografických procesech se regionální kontrasty nejvíce prohlubují.

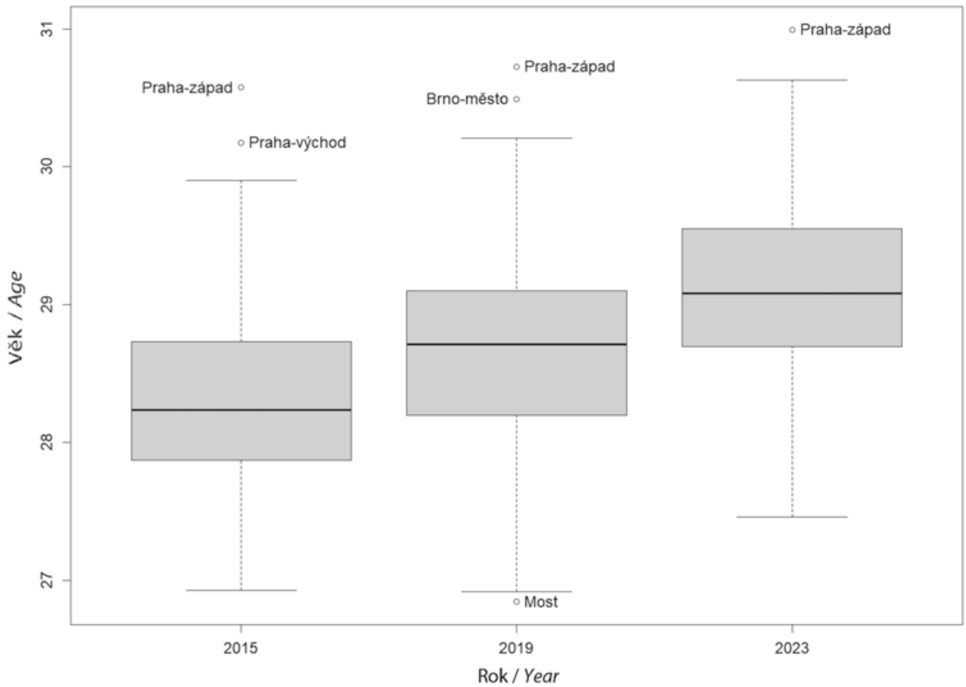
Nejprve zde na začátku analytických částí článku, je nicméně vhodné podpořit konkrétními čísly nemožnost zahrnutí Prahy do regionálního srovnání. Pro tento účel je využita Mahalanobisova vzdálenost.

Jediným okresem, který byl ve všech sledovaných letech dle Mahalanobisovy vzdálenosti označen za vícerozměrné odlehlé pozorování, je Hlavní město Praha. Zároveň je jeho odlehlost patrná i z komparace s jinými odlehlými okresy. Z toho důvodu se tento okres v analýze (deskriptivní analýze ani shlukové analýze) nadále již nevyskytuje tak, aby nezkresloval výsledky a současně byly všechny provedené analýzy vzájemně porovnatelné. Pro možnost komparace Prahy s výsledky ostatních okresů, v případě zájmu čtenáře, jsou následně konkrétní hodnoty za Prahu zařazeny do přílohy 1.

Z grafu 1 vyplývá, že mezi lety 2015 a 2019 došlo v okresech Česka k nárůstu úhrnné plodnosti, v roce 2023 naopak došlo ke znatelnému poklesu. Zajímavým výsledkem této analýzy je skutečnost, že v roce 2019 byla naměřena minimální úhrnná plodnost (v okresu Plzeň-sever) vyšší než mediánová úhrnná plodnost pro všechny okresy ČR v roce 2023. Úhrnná plodnost ve zmíněném okresu byla dokonce i vyšší než třetí kvartil úhrnné plodnosti z roku 2023. Maximální zjištěná hodnota úhrnné plodnosti v roce 2023 (v okresu Kroměříž) byla nižší než první kvartil sledující stejnou charakteristiku v roce 2019. Přestože se naměřené hodnoty úhrnné plodnosti v čase výrazně mění, diference mezi okresy, charakterizovaná mezikvartilovým rozpětím, se ve všech sledovaných letech pohybovala kolem jedné desetiny, jak je patrné z krabičkových grafů.

Ve třech okresech v roce 2023 úhrnná plodnost nepřekročila ani hranici 1,3, jedná se o okresy Cheb, Tachov a Karlovy Vary (okres Cheb geograficky sousedí se zbylými jmenovanými okresy a společně s Karlovými Vary se nachází v Karlovarském kraji, okres Tachov je součástí Plzeňského kraje). Hodnoty těchto územních celků jsou považovány za odlehlé.

Graf 2: Regionální diferenciace průměrného věku matky při narození prvního živého dítěte v letech 2015, 2019 a 2023 (soubor okresů bez Prahy) – krabičkový graf / Regional differentiation of the mean age of mothers at first live birth in 2015, 2019, and 2023 (a set of districts without Prague) – box plot



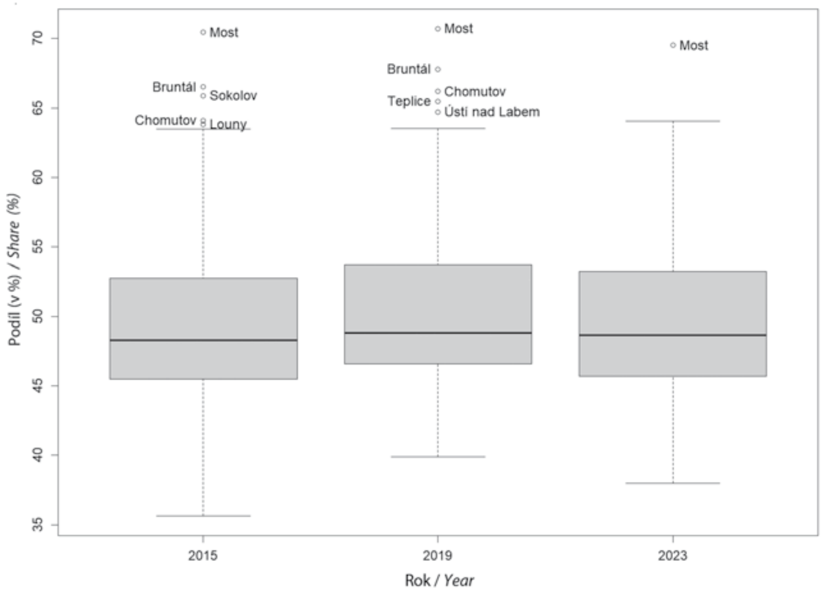
Zdroj: Vlastní zpracování na základě dat ČSÚ.
Source: *Authors' calculation based on CZSO data.*

Ve všech sledovaných letech na grafu 2 byl zaznamenán nejvyšší průměrný věk prvorodiček v okresu Praha-západ. Zároveň ve všech letech údaj figuroval jako odlehlé pozorování. V roce 2015 byl také odlehlým pozorováním okres Praha-východ, v roce 2019 Brno-město. Ve všech uvedených případech byly okresy označeny za odlehlá pozorování z důvodu příliš vysokého průměrného věku prvorodiček, což poukazuje na odlišný životní styl a přístup k plánování rodinného života žen. V roce 2019 byl kromě již výše zmíněných okresů za odlehlé pozorování považován taktéž okres Most, kde však průměrný věk prvorodiček byl považován za odlehlé pozorování z důvodu nízké hodnoty sledovaného ukazatele. Také v roce 2015 byl v tomto okresu v rámci celorepublikového srovnání nejnižší průměrný věk prvorodiček. Naopak v roce 2023 okres Most dosahoval až sedmé nejnižší hodnoty. Nejnižší a druhý nejnižší průměrný věk prvorodiček byl v okresech Děčín a Chomutov, tedy v okresech z Ústeckého kraje. V průběhu času lze na základě všech tří kvartilů konstatovat narůstající průměrný věk prvorodiček.

Z grafu 3 lze vyčíst, že v roce 2015 bylo v porovnání s dalšími sledovanými roky napozorováno nejvyšší variační rozpětí podílu živě narozených dětí mimo manželství. Rozdíl mezi nejvyšší a nejnižší sledovanou hodnotou činil v tomto roce při zohledňování i odlehlých hodnot po zaokrouhlení 34,83 procentního bodu. Je zajímavé, že na rozdíl od předchozích sledovaných ukazatelů je zde poměrně stabilní trend, kdy medián ani kvartily se v čase v podstatě nemění.

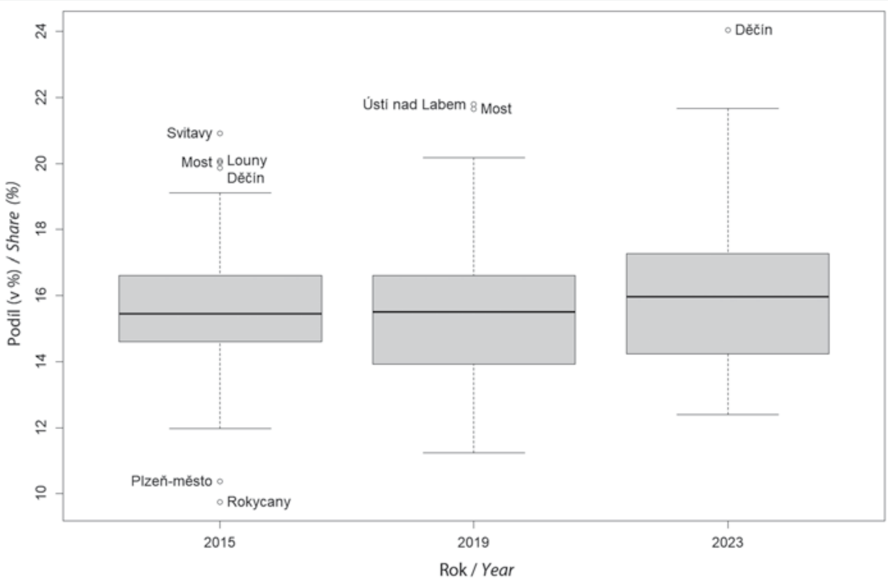
V okresu Most se pravidelně rodilo nejvyšší procento dětí mimo manželský svazek, přičemž v letech 2015 a 2019 byla hodnota vyšší než 70 % (v roce 2023 69,5 %). Nejenom v okresu Most, ale i v jiných okresech Ústeckého kraje se narodil výrazný podíl dětí mimo manželství. Naopak okres Brno-venkov patřil během všech tří let mezi tři okresy s nejnižším procentem živě narozených dětí mimo manželství.

Graf 3: Regionální diferenciace podílu živě narozených dětí mimo manželství v letech 2015, 2019 a 2023 (soubor okresů bez Prahy) – krabičkový graf / Regional differentiation of the percentage of live births outside marriage in 2015, 2019, and 2023 (a set of districts without Prague) – box plot



Zdroj: Vlastní zpracování na základě dat ČSÚ.
Source: *Authors' calculation based on CZSO data.*

Graf 4: Regionální diferenciace podílu živě narozených dětí třetího a vyššího pořadí v letech 2015, 2019 a 2023 (soubor okresů bez Prahy) – krabičkový graf / Regional differentiation of the percentage of third- and higher-order live births in 2015, 2019, and 2023 (a set of districts without Prague) – box plot



Zdroj: Vlastní zpracování na základě dat ČSÚ.
Source: *Authors' calculation based on CZSO data.*

V grafu 4 je patrný nárůst mediánu podílu narozených dětí třetího a vyššího pořadí. Vrátíme-li se ke grafu 1, medián úhrnné plodnosti v roce 2023 byl ze tří sledovaných let nejnižší, medián třetího a vyššího pořadí byl však za tři sledované roky nejvyšší, tím pádem lze konstatovat, že došlo z relativního hlediska primárně k poklesu plodnosti potenciálních prvorodiček a matek, které by porodily druhé dítě, a v menší míře matek, které již živě porodily alespoň 2 děti.

V roce 2015 bylo, v porovnání s následujícími sledovanými roky, zaznamenáno nejnižší mezikvartilové rozpětí. Odlehlé hodnoty však způsobily, že variační rozpětí v tomto roce bylo naopak oproti rokům 2019 a 2023 nejvyšší. Za cenu určitého zjednodušení lze konstatovat, že diferenciace okresů z hlediska plodnosti třetího a vyššího pořadí ve sledovaném období výrazně vzrostla.

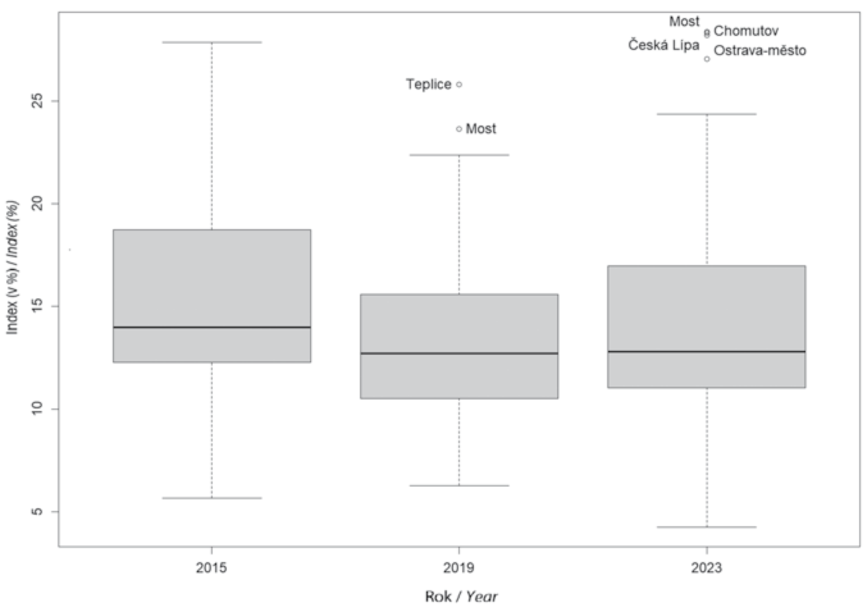
Výrazný podíl dětí narozených ve třetím a vyšším pořadí byl opět zaznamenán primárně v okresech Ústeckého kraje, v roce 2015 naopak dva okresy z Plzeňského kraje dosáhly v porovnání se zbylými sledovanými okresy příliš nízkých hodnot, a tak byly označeny za odlehlá pozorování.

Taktéž tento ukazatel poukazuje na skutečnost, že vysokých (odlehlých) hodnot nabývaly okresy primárně z Ústeckého kraje, okres Most byl dokonce jako jediný v letech 2019 a zároveň 2023 odlehlým pozorováním. Největší variační rozpětí bylo zaznamenáno v roce 2023, kdy v okrese Most tento ukazatel nabýval hodnoty 28,37 %, zatímco v okresu Klatovy 4,27 %.

Pro sledování variability napříč různými proměnnými i různým časovým obdobím byl doplněn výpočet variačních koeficientů.

Nejvyšší variabilita byla z hlediska typu proměnných napozorována u indexu UPT. Naopak nejnižší hodnoty variačního koeficientu byly napozorovány pro proměnnou průměrný věk prvorodiček, a to i přesto, že v rámci proměnné úhrnná plodnost se v datovém souboru vyskytuje méně odlehlých pozorování. Tato proměnná se od ostatních liší tím, že nejvyšší variabilita byla napozorována v roce 2019. Variabilita podílu živě narozených dětí mimo manželství v čase klesá.

Graf 5: Regionální diferenciace indexu UPT v letech 2015, 2019 a 2023 (soubor okresů bez Prahy) – krabičkový graf / Regional differentiation of the index of induced abortions for non-medical reasons per number of live births in 2015, 2019, and 2023 (a set of districts without Prague) – box plot



Zdroj: Vlastní zpracování na základě dat ČSÚ.
Source: *Authors' calculation based on CZSO data.*

Tab. 2: Vývoj variačních koeficientů (v %) v letech 2015, 2019 a 2023 (soubor okresů bez Prahy)
Change in the coefficients of variation (in %) in 2015, 2019, and 2023 (a set of districts excluding Prague)

Ukazatel / *Indicator*	**Rok / *Year***		
	2015	**2019**	**2023**
Podíl živě narozených dětí mimo manželství (v %) *Share of live births outside marriage (%)*	14,84	13,59	13,73
Úhrnná plodnost / *Total fertility rate*	5,34	4,49	5,40
Podíl živě narozených třetího a vyššího pořadí (v %) *Share of third- and higher-order live births (%)*	13,53	14,42	15,10
Index UPT / *Index of induced abortions*	29,58	30,66	36,24
Průměrný věk matky při narození prvního živého dítěte *Average age of mother at first live birth*	2,62	2,66	2,33

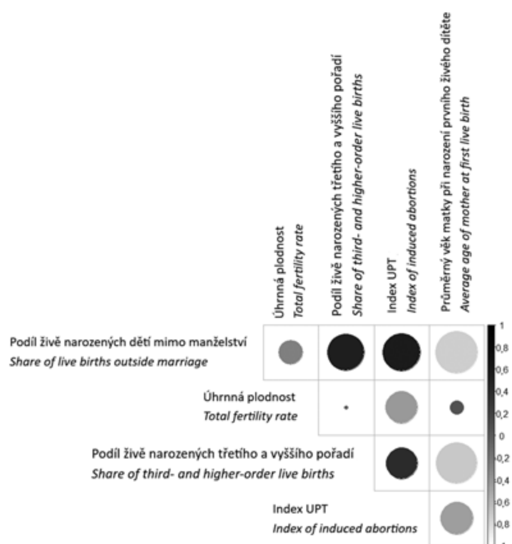
Zdroj: Vlastní zpracování na základě dat ČSÚ.
Source: *Authors' calculation based on CZSO data.*

Pro zbylé proměnné platí, že variační koeficient nabývá nejvyšších hodnot pro data z roku 2023.

Pro doplnění deskriptivního pohledu byly také pro každý rok (2015, 2019, 2023) vytvořeny korelační matice, které porovnávají závislost mezi sledovanými proměnnými. Vzhledem ke skutečnosti, že během výše zmíněných let nedošlo k výrazným a interpretačně příliš zajímavým změnám v rámci hodnot monotónních závislostí, obsahuje tato část pouze grafický výstup korelační matice za rok 2023. Korelační matice vyobrazující závislosti v letech 2015 a 2019 byly zařazeny jako přílohy 2 a 3. Pro provedení výpočtů byl využit Spearmanův korelační koeficient.

Hodnota Spearmanova korelačního koeficientu nabývá hodnot od –1 do 1. Čím větší je kruh mezi dvěma pozorovanými proměnnými, tím větší je mezi proměnnými závislost, tmavá barva kruhu znázorňuje pozitivní závislost (korelační koeficient nabývá kladné hodnoty), světle šedý kruh naopak vyobrazuje negativní závislost (korelační koeficient nabývá záporné hodnoty). Malý šedý kruh znázorňuje nízkou (téměř nulovou) závislost.

Graf 6: Korelační matice sledující závislost mezi sledovanými proměnnými v roce 2023
Correlation matrix monitoring the dependence between selected variables in 2023



Zdroj: Vlastní zpracování na základě dat ČSÚ.
Source: *Authors' calculation based on CZSO data.*

Z grafu 6 je patrných několik poměrně očekávaných závislostí, jako je při vyšším procentu mimomanželsky narozených nižší průměrný věk prvorodiček a podobně, jak představí následný výčet.

Například mezi okresy existuje silná pozitivní závislost mezi proměnnými:

- Procento živě narozených dětí mimo manželství a index uměle přerušených těhotenství nevynucených zdravotními důvody na počet živě narozených (0,594), jelikož obě proměnné poukazují na liberálnější postoje k manželství a rodičovství i patrnou větší nestabilitu partnerství v regionu.
- Procento narozených třetího a vyššího pořadí a procento živě narozených dětí mimo manželství (0,555), kdy v některých okresech může být rozšířenější kohabitace i při vyšším počtu dětí v rodině, opět to může souviset se socioekonomickým statusem části populace.
- Procento narozených třetího a vyššího pořadí a index UPT (0,413), tento vztah poukazuje na výskyt oblastí, ve kterých je běžnější mít více dětí, a kde se zároveň častěji objevují situace, kdy je těhotenství neplánované nebo časově nevhodné, což může vést k vyšší míře záměrně přerušených koncepcí.

Naopak silná negativní závislost se vyskytuje mezi proměnnými:

- Průměrný věk matek při narození prvního živého dítěte a procento živě narozených dětí mimo manželství (–0,737), platí, že mladší ženy častěji mají děti v kohabitaci, nebo ještě před vstupem do manželství. V regionech s nižším věkem při narození prvního dítěte je proto běžnější rodičovství mimo manželství, zatímco v oblastech s odkládaným vstupem do mateřství bývá rodičovství častěji spojeno s manželským svazkem.
- Průměrný věk matek při narození prvního živého dítěte a procento narozených třetího a vyššího pořadí (–0,703). V regionech, kde ženy vstupují do mateřství v mladším věku, je obvykle více času i prostoru pro pokračování reprodukce, což vede k vyššímu podílu dětí vyššího pořadí. Naopak tam, kde dochází k výraznému odkládání prvního porodu do vyššího věku, se snižuje pravděpodobnost dosažení tří a více dětí, což se promítá do nižšího podílu porodů vyššího pořadí.
- Úhrnná plodnost a index UPT (–0,429), z logické podstaty věci proměnných je patrné, že signifikantní poměr chtěných uměle přerušených těhotenství a počtu živě narozených negativně ovlivňuje úhrnnou plodnost.

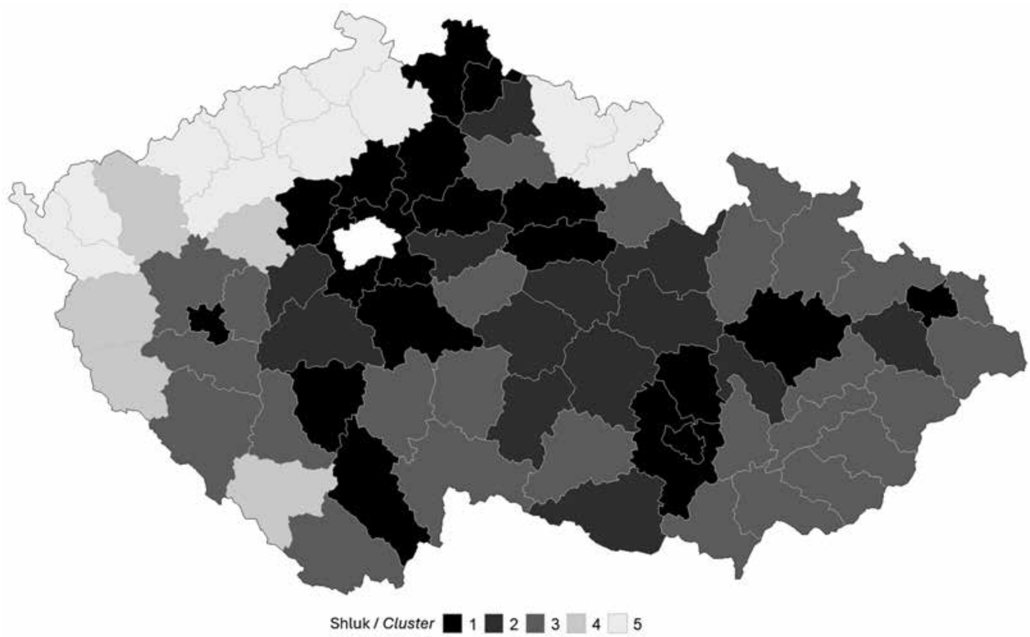
SHLUKOVÁ ANALÝZA

Regionální rozdíly v reprodukčním chování okresů jsou výsledkem dlouhodobě působících sociálních, kulturních i ekonomických procesů, které vytvářejí specifické prostorové vzorce. Zatímco deskriptivní část ukázala, jak se jednotlivé ukazatele v okresech liší, tato kapitola se zabývá otázkou, zda lze na základě kombinace sledovaných charakteristik rozpoznat skupiny okresů s obdobným reprodukčním profilem. Tato strukturace českého prostoru umožňuje zachytit vývoj komplexních vztahů mezi ukazateli v letech 2015, 2019 a 2023, které nejsou patrné při izolovaném pohledu na jednotlivé charakteristiky.

Shluková analýza je zde využita jako nástroj k odhalení těchto vzorců a výsledné shluky okresů poskytují ucelený obraz o tom, jak se různé podoby reprodukčního chování v českých regionech koncentrují, které oblasti se sobě nejvíce podobají a v čem se naopak odlišují. Kapitola tak přináší interpretaci regionálních typů reprodukčního chování a ukazuje, jak stabilní nebo proměnlivé tyto typy jsou. S ohledem na srozumitelnou interpretaci i dobrou čitelnost v černobílém zobrazení je v kartogramu zobrazena varianta s pěti shluky.

Z obrázku 1 lze na první pohled rozklíčovat patrné zákonitosti. Shlukovou analýzou byly rozděleny okresy do celkem 5 shluků. Je zřejmé, že obdobným způsobem se proměnné ve sledovaných letech vyvíjely v příměstských regionech (okolí Prahy a Brna), v okresu Plzeň-město a dalších regionech s významným vlivem městského centra (shluk 1). Pro tyto okresy je typické odkládání založení rodin (nejvyšší průměrný věk matky při narození prvního živého dítěte), což může také mít vliv na skutečnost, že pouze nízké procento dětí se zde narodí ve 3. a vyšším pořadí. Tyto děti se primárně rodí sezdaným párům. Obecně by se dal tento shluk z hlediska reprodukčního chování pojmout jako okresy, kde rodiče zakládají stabilní

Obrázek 1: Kartogram zobrazující shlukování okresů Česka (bez Prahy) do 5 shluků v letech 2015, 2019 a 2023
Choropleth map showing the clustering of Czech districts (excluding Prague) into five clusters in 2015, 2019. and 2023



Zdroj: Vlastní zpracování na základě dat ČSÚ.
Source: *Authors' calculation based on CZSO data.*

rodiny ve vyšším věku. Jedná se o poměrně početný shluk, který je tvořen 19 okresy.

Jiné reprodukční chování bylo zjištěno na přelomu Čech a Moravy (shluk 2). Pro tuto oblast je typická nejvyšší úhrnná plodnost, která je podporována nejnižším indexem UPT, tedy nejnižším poměrem UPT, o kterých nebylo rozhodnuto na základě zdravotního stavu, a počtem živě narozených dětí. Do druhého shluku je seskupeno 13 okresů.

Typická je také oblast hraničního východního přechodu se Slovenskem, tedy oblast Moravy a Slezska, zároveň byl obdobný vývoj charakteristik sledován v jižní a jihozápadní oblasti republiky (shluk 3). Tato oblast nabývá z hlediska celorepublikového srovnání v rámci sledovaných proměnných primárně průměrných hodnot, v ničem nevybočuje. Jedná se o nejpočetnější shluk vytvořený touto analýzou, obsahuje 27 okresů.

Obdobné vzorce chování byly napozorovány rovněž v Ústeckém a Karlovarském kraji (shluk 5), jediným okresem z těchto dvou krajů, který se od zbylých lišil, byl okres Karlovy Vary (shluk 4). Součástí stejného shluku jako je většina okresů ze Severozápadu jsou také okresy Trutnov a Náchod. Pro tyto okresy je typický vysoký podíl dětí narozených nesezdaným párům, nejnižší průměrný věk matky při narození prvního živého dítěte, nejvyšší index UPT, což souvisí s nižšími hodnotami ukazatele úhrnné plodnosti, či nejvyšší podíl dětí narozených ve třetím a vyšším pořadí. Tento shluk působí jako kontrast shluku 1, kdy na rozdíl od prvního zmiňovaného shluku jsou zde zakládány rodiny sice mladšími lidmi (resp. mladšími prvorodičkami), nicméně samotné početí a jeho budoucnost můžou a často i jsou nestabilní, o čemž svědčí vysoký podíl dětí narozených nesezdaným párům a vysoký index UPT. Shluk je tvořen 12 okresy.

Nejméně početný shluk (shluk 4) figuruje jako určitý mezičlánek mezi shluky 3 a 5, a to nejenom z hlediska geografické polohy okresů, ale také z hlediska hodnot sledovaných proměnných. Tento shluk obsahuje pouze 5 okresů.

Výsledky klasifikace okresů je možné dále podrobněji popsat opět pomocí deskriptivní statistiky.

Tab. 3: Průměrné hodnoty a variační koeficienty sledovaných proměnných napříč shluky v roce 2015
Mean values and coefficients of variation of the monitored variables across clusters in 2015

Shluk *Cluster*	**Podíl živě narozených dětí mimo manželství (v %)** *Share of live births outside marriage (%)*		**Úhrnná plodnost** *Total fertility rate*		**Podíl živě narozených třetího a vyššího pořadí (v %)** *Share of third- and higher-order live births (%)*		**Index UPT (v %)** *Index of induced abortions (%)*		**Průměrný věk matky při narození prvního živého dítěte** *Average age of mother at first live birth*	
	Průměr *Mean*	**Var. koef. (v %)** *Var. coeff. (%)*	**Průměr** *Mean*	**Var. koef. (v %)** *Var. coeff. (%)*	**Průměr** *Mean*	**Var. koef. (v %)** *Var. coeff. (%)*	**Průměr** *Mean*	**Var. koef. (v %)** *Var. coeff. (%)*	**Průměr** *Mean*	**Var. koef. (v %)** *Var. coeff. (%)*
1	46,17	11,68	1,61	3,71	13,76	11,06	14,01	23,56	29,01	2,27
2	46,72	11,55	1,61	3,45	16,08	9,41	12,32	26,73	28,39	2,29
3	48,28	11,03	1,54	3,62	15,27	9,97	14,48	20,06	28,28	2,28
4	54,76	9,49	1,59	3,48	15,72	9,43	15,20	17,70	27,69	2,31
5	60,97	8,53	1,57	3,37	18,21	7,82	22,59	11,80	27,45	2,36

Zdroj: Vlastní zpracování na základě dat ČSÚ.
Source: *Authors' calculation based on CZSO data.*

Následující oddíl detailně popisuje vývoj sledovaných ukazatelů ve stanovených shlucích ve sledovaných letech.

V roce 2015 byl, jak ukazuje tabulka 3, napříč všemi shluky zaznamenán nejnižší průměrný věk matky při narození prvního živého dítěte, tento ukazatel ve všech shlucích v čase nabýval vyšších hodnot, což potvrzuje zjištění patrné z grafu 2. Nejvíce heterogenní z hlediska porovnání variačních koeficientů je shluk 1, což je primárně dáno vysokým počtem v něm sdružených okresů. Naopak nejvíce homogenní jsou okresy ve shluku 5.

Z tabulky 4 v porovnání s předchozím sledovaným obdobím, s rokem 2015, došlo kromě nárůstu průměrného věku prvorodičky také k výraznému nárůstu úhrnné plodnosti (což navazuje na komentář grafu 1) a zároveň téměř ve všech shlucích (kromě shluku 4) došlo k poklesu indexu UPT a nárůstu podílu živě narozených dětí mimo manželství. Okresy členěné do shluku 2 jsou v porovnání s předchozím

Tab. 4: Průměrné hodnoty a variační koeficienty sledovaných proměnných napříč shluky v roce 2019
Mean values and coefficients of variation of the monitored variables across clusters in 2019

Shluk *Cluster*	**Podíl živě narozených dětí mimo manželství (v %)** *Share of live births outside marriage (%)*		**Úhrnná plodnost** *Total fertility rate*		**Podíl živě narozených třetího a vyššího pořadí (v %)** *Share of third- and higher-order live births*		**Index UPT (v %)** *Index of induced abortions (%)*		**Průměrný věk matky při narození prvního živého dítěte** *Average age of mother at first live birth*	
	Průměr *Mean*	**Var. koef. (v %)** *Var. coeff. (%)*	**Průměr** *Mean*	**Var. koef. (v %)** *Var. coeff. (%)*	**Průměr** *Mean*	**Var. koef. (v %)** *Var. coeff. (%)*	**Průměr** *Mean*	**Var. koef. (v %)** *Var. coeff. (%)*	**Průměr** *Mean*	**Var. koef. (v %)** *Var. coeff. (%)*
1	46,73	9,82	1,75	3,30	13,96	11,19	11,86	22,32	29,42	1,93
2	47,49	5,20	1,81	3,65	16,00	8,29	11,46	18,83	28,71	1,64
3	49,70	11,95	1,69	4,18	15,24	12,61	11,87	22,37	28,56	1,93
4	53,18	5,98	1,73	4,66	15,39	6,61	17,29	19,32	28,36	0,95
5	61,21	8,11	1,72	3,34	18,49	11,30	20,02	13,84	27,72	2,10

Zdroj: Vlastní zpracování na základě dat ČSÚ.
Source: *Authors' calculation based on CZSO data.*

Tab. 5: Průměrné hodnoty a variační koeficienty sledovaných proměnných napříč shluky v roce 2023
Mean values and coefficients of variation of the monitored variables across clusters in 2023

Shluk *Cluster*	**Podíl živě narozených dětí mimo manželství (v %)** *Share of live births outside marriage (%)*		**Úhrnná plodnost** *Total fertility rate*		***Podíl živě narozených třetího a vyššího pořadí (v %)*** *Share of third- and higher-order live births (%)*		**Index UPT (v %)** *Index of induced abortions (%)*		**Průměrný věk matky při narození prvního živého dítěte** *Average age of mother at first live birth*	
	Průměr *Mean*	**Var. koef. (v %)** *Var. coeff. (%)*	**Průměr** *Mean*	**Var. koef. (v %)** *Var. coeff. (%)*	**Průměr** *Mean*	**Var. koef. (v %)** *Var. coeff. (%)*	**Průměr** *Mean*	**Var. koef. (v %)** *Var. coeff. (%)*	**Průměr** *Mean*	**Var. koef. (v %)** *Var. coeff. (%)*
1	45,49	9,07	1,46	3,56	14,57	9,37	14,63	30,23	29,71	1,88
2	47,52	7,20	1,56	2,89	16,57	7,58	11,87	20,53	29,00	1,18
3	48,92	11,68	1,50	3,97	15,09	11,51	11,87	25,77	29,13	1,23
4	51,00	4,78	1,39	7,30	17,28	7,36	13,57	31,21	28,90	1,97
5	61,07	6,75	1,43	5,62	19,92	10,03	22,33	20,11	28,08	1,66

Zdroj: Vlastní zpracování na základě dat ČSÚ.
Source: *Authors' calculation based on CZSO data.*

obdobím více homogenní. Zároveň pro shluk 1 již neplatí, že by okresy jej tvořící byly nejvíce heterogenní, nejvyšší variabilita napříč sledovanými proměnnými byla napozorována v nejpočetnějším shluku 3. Naopak nejvíce homogenním shlukem pro tento rok byl nejméně početný shluk 4.

Z tabulky 5 je patrné, že v porovnání s rokem 2019 opět došlo ve všech shlucích v průměru k nárůstu průměrného věku matky při narození prvního živého dítěte a k poklesu úhrnné plodnosti, a i přes tento pokles, kromě shluku 3 došlo k nárůstu podílu dětí narozených ve třetím a vyšším pořadí, což poukazuje na nižší tendenci zakládání rodin či rozšíření rodin s 1 dítětem na rodiny se 2 dětmi. Nejvíce homogenní v tomto roce byly okresy tvořící shluk 2. V porovnání se zbylými shluky jsou o něco více heterogenní shluky 3 a 4.

ZÁVĚR

Analýza reprodukčního chování okresů Česka v letech 2015, 2019 a 2023 ukázala, že prostorové rozdíly mezi regiony jsou výrazné a přetrvávají napříč sledovaným obdobím. Deskriptivní statistika potvrdila, že sledované ukazatele se mezi okresy liší nejen úrovní, ale i rozsahem variability. Analýza ukázala, že zatímco distribuce úhrnné plodnosti vykázala mezi roky 2019 a 2023 výrazný pokles v naměřených hodnotách, průměrný věk matky při narození prvního dítěte dlouhodobě roste. Poměrně výrazná variabilita byla napozorována u indexu UPT, mimomanželská plodnost zůstávala z hlediska měr polohy poměrně konstantní, a to i přes častý výskyt extrémních hodnot v Severozápadním regionu Česka.

Shluková analýza rozdělila okresy do pěti skupin s odlišnými reprodukčními profily. Jeden shluk zahrnuje okresy s nízkým průměrným věkem matek při prvním narození živého dítěte, vysokým podílem mimomanželsky narozených a vyšším indexem UPT. Další shluk tvoří okresy s opačnými charakteristikami – vyšším věkem prvorodiček a nízkými hodnotami ostatních ukazatelů. Třetí shluk sdružuje okresy s průměrnými hodnotami většiny ukazatelů, čtvrtý shluk se vyznačuje vyšším zastoupením dětí třetího a vyššího pořadí a pátý shluk zahrnuje okresy se specificky vyšší intenzitou UPT. Tato struktura potvrzuje, že české okresy lze rozdělit do několika jasně odlišitelných typů reprodukčního chování, které nejsou dány pouze geografickou polohou, ale kombinací více demografických charakteristik.

Stanovené cíle a související výzkumné otázky byly ve všech sledovaných krocích naplněny a zodpovězeny. První cíl byl splněn detailním popisem regionální diferenciace reprodukčního chování na základě pěti klíčových ukazatelů, což potvrdilo výraznou a dlouhodobě přítomnou prostorovou variabilitu. Druhý cíl byl naplněn porovnáním vývoje těchto ukazatelů mezi lety 2015, 2019 a 2023, komparace odhalila jak stabilní rysy, tak dílčí změny ve variabilitě ukazatelů. Třetí cíl byl splněn identifikací pěti shluků okresů s odlišnými reprodukčními profily, které poskytují ucelený obraz o tom, jak se jednotlivé kombinace ukazatelů v českém prostoru sdružují.

Výsledky studie tak přinášejí komplexní a aktuální pohled na prostorovou strukturu reprodukčního chování v Česku a vytvářejí základ pro interpretace jeho socioekonomických a kulturních souvislostí.

Literatura

- Bakuncová, K., – Löster, T. 2023. Time Series Clustering Of Fertility in European Union. *International Days of Statistics and Economics*, s. 19–35. Praha: VŠE. https://msed.vse.cz/msed_2023/sbornik/toc.html.
- Bartoňová, D. 1999. Vývoj regionální diferenciace věkové struktury se zřetelem k územním rozdílům ve vývoji reprodukce v České republice. *Geografie – Sborník ČGS*, 104(1), s. 13–23. ISSN 1212-0014. https://doi.org/10.37040/geografie1999104010013.
- Bartoňová, D. 2001. Demografické chování populace České republiky v regionálním a evropském kontextu. In Hampl, M. (Ed.) *Regionální vývoj: specifika české transformace, evropská integrace a obecná teorie*, s. 45–73, Praha: DemoArt.
- ČSÚ. 2025. *Narození*. https://csu.gov.cz/narozeni.
- Fialová, L. 2018. *Sociologická encyklopedie*. Dostupné z: https://encyklopedie.soc.cas.cz/w/Regionalizace.
- Hindls, R. – Hronová, S. – Seger, J. – Fischer, J. 2007. *Statistika pro ekonomy*. Praha: Professional Publishing. 415 s. ISBN 978-80-86946-43-6.
- Hon, F. 2020. Současná úroveň plodnosti třetího pořadí v České republice. In *Sborník příspěvků z konference RELIK*, s. 129–139. Praha: VŠE. https://relik.vse.cz/2020/download/pdf/363-Hon-Filip-paper.pdf.
- Kocourková, J. – Šťastná, A., Idlbeková, B. – Slabá, J. 2023. Trendy plodnosti a potratovosti v Česku. *Časopis lékařů českých*, 162, s. 299–306. https://www.prolekare.cz/casopisy/casopis-lekaru-ceskych/2023-7-8-1/trendy-plodnosti-a-potratovosti-v-cesku-136675.
- Kurkin, R. 2015. *Faktory plodnosti a regionální diferenciace plodnosti v České republice po roce 1991* [Disertační práce, Univerzita Karlova]. https://explorer.cuni.cz/publication/502597.
- Lacko, J. 2024. *RCzechia: Spatial Objects of the Czech Republic*. https://rczechia.jla-data.net. https://doi.org/10.21105/joss.05082.
- Microsoft Corporation. 2024. *Microsoft Excel* [Software]. https://office.microsoft.com/excel.
- Paloncyová, J. – Hašková, H. – Höhne, S. – Kuchařová, V. – Žáčková, L., – Lukić, A. 2021. *Bezdětní a postoje k rodičovství.* VÚPSV. 170 s. https://katalog.vupsv.cz/fulltext/vz_487.pdf.
- Pebesma, E. 2018. Simple Features for R: Standardized Support for Spatial Vector Data. *The R Journal*, 10(1), s. 439–446. https://doi.org/10.32614/RJ-2018-009.
- Pecáková, I. 2011. *Statistika v terénních průzkumech*. Praha: Professional Publishing. 236 s. ISBN 978-80-7431-039-3.
- R Core Team. 2024. *R: A language and environment for statistical computing* [Software]. R Foundation for Statistical Computing. https://www.R-project.org/.
- Rychtaříková, J. 2007. Regionální diferenciace plodnosti v průřezové a kohortní perspektivě. In Kučera, T. – Polášek, V. (Eds.) *Sborník příspěvků XXXVII. Výroční demografické konference ČDS*, s. 92–103, Olomouc: ČDS.
- Řezanková, H. 2024. *Statistické metody se zaměřením na kategoriální data*. Praha: Oeconomica. 210 s. https://oeconomica.vse.cz/publikace/statisticke-metody-se-zamerenim-na-kategorialni-data-e-kniha/.
- Slabá, J. – Kocourková, J. – Šťastná, A. 2024. The fertility timing gap: the intended and real timing of childbirth. *Journal of Biosocial Science*, 56(3), s. 504–517. doi:10.1017/S002193202400004X.
- SPSS Inc. 2009. *PASW Statistics for Windows, Version 18.0* [Software]. SPSS Inc.
- Stípková, M. 2015. Ideational and Economic Causes of the Rise in Non-marital Childbearing in the Czech Republic. *European Journal of Population*, 31(5), s. 473–494. https://doi.org/10.1007/s10680-015-9350-4.
- Šídlo, L. 2008. Faktory ovlivňující regionální diferenciaci plodnosti v Česku na počátku 21. století. *Demografie, revue pro výzkum populačního vývoje,* 50(3), s. 186–198.

- Šídlo, L. – Šprocha, B. 2018. Odkládání mateřství a regionální diferenciace plodnosti v Česku a na Slovensku. *Geografie*, 123(3), s. 407–436. https://doi.org/10.37040/geografie2018123030407.
- Šimpach, O. – Šimpachová Pechrová, M. 2016. Cluster Analysis Of EU's Regions According to Demographics Criteria. *International Days of Statistics and Economics*, s. 1806–1815. Praha: VŠE. https://msed.vse.cz/msed_2016/article/166-Simpach-Ondrej-paper.pdf.
- Šprocha, B. – Šídlo, L. 2016a. Regionálne rozdiely v charaktere plodnosti v Česku a na Slovensku. *Sborník příspěvků z konference RELIK*, s. 559–571. Dostupné z: https://relik.vse.cz/2016/download/pdf/57-Sprocha-Branislav-paper.pdf.
- Šprocha, B. – Šídlo, L. 2016b. Spatial Differentiation and Fertility Postponement Transformation in Czechia. *AUC Geographica*. Dostupné z: https://karolinum.cz/data/clanek/3189/Geogr_2_2016_08_Sprocha.pdf. https://doi.org/10.14712/23361980.2016.18.
- Waldaufová, E. – Šťastná, A. 2025. Mezi ideálem a realitou: současný pokles porodnosti očima české populace. *Demografie*, *revue pro výzkum populačního vývoje*, 67(2), s. 83–101. https://doi.org/10.54694/dem.0359.
- Wei, T. – Simko, V. 2024. *R package „corrplot": Visualization of a Correlation Matrix*. https://github.com/taiyun/corrplot.
- Wickham, H. – Bryan, J. 2025. *readxl: Read Excel Files.* https://CRAN.R-project.org/package=readxl.
- Wickham, H. – François, R. – Henry, L. – Müller, K. – Vaughan, D. 2023. *dplyr: A Grammar of Data Manipulation* [Software]. R package version 1.1.4. https://CRAN.R-project.org/package=dplyr.
- Wickham, H. – Chang, W. – Henry, L. – Pedersen, T. L. – Takahashi, K. – Wilke, C. – Woo, K. – Yutani, H. – Dunnington, D. – Brand, T. van den. 2025. *ggplot2: Create Elegant Data Visualisations Using the Grammar of Graphics*. https://ggplot2.tidyverse.org.
- Wickham, H. – Vaughan, D. – Girlich, M. 2024. *tidyr: Tidy Messy Data*. https://tidyr.tidyverse.org.

KRISTÝNA DVOŘÁKOVÁ

Je studentkou magisterského programu Ekonomická demografie na Fakultě informatiky a statistiky VŠE v Praze, dále působí jako výzkumný pracovník na Ústavu informatiky AV ČR. Věnuje se kvantitativním metodám a modelování demografických procesů.

FILIP HON

Je absolventem Fakulty informatiky a statistiky VŠE v Praze. Od roku 2018 je členem Katedry demografe VŠE v Praze a výzkumným pracovníkem RILSA v Praze. Zabývá se demografickým prognózováním, výzkumem plodnosti a sociální ochranou.

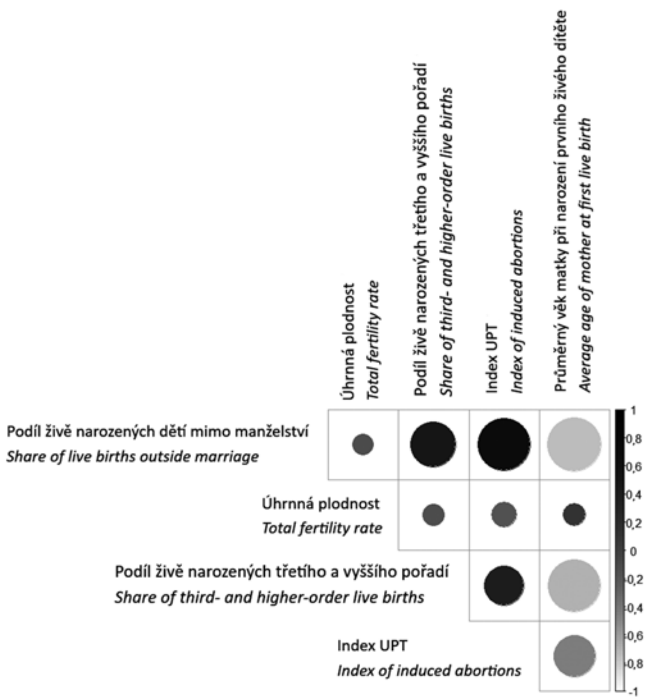
PŘÍLOHA / APPENDIX

Příloha 1: Vývoj sledovaných ukazatelů v letech 2015, 2019 a 2023 pro Prahu
Changes in the selected indicators in 2015, 2019, and 2023 for the district of the Capital City of Prague

Ukazatel / *Indicator*	**Rok / *Year***		
	2015	**2019**	**2023**
Podíl živě narozených dětí mimo manželství (v %) *Share of live births outside marriage (%)*	61,08	56,35	60,39
Úhrnná plodnost / *Total fertility rate*	1,50	1,55	1,33
Podíl živě narozených třetího a vyššího pořadí (v %) *Share of third- and higher-order live births (%)*	10,73	10,42	9,82
Index UPT / *Index of induced abortions*	13,14	10,46	11,45
Průměrný věk matky při narození prvního živého dítěte *Average age of mother at first live birth*	30,09	30,24	30,58

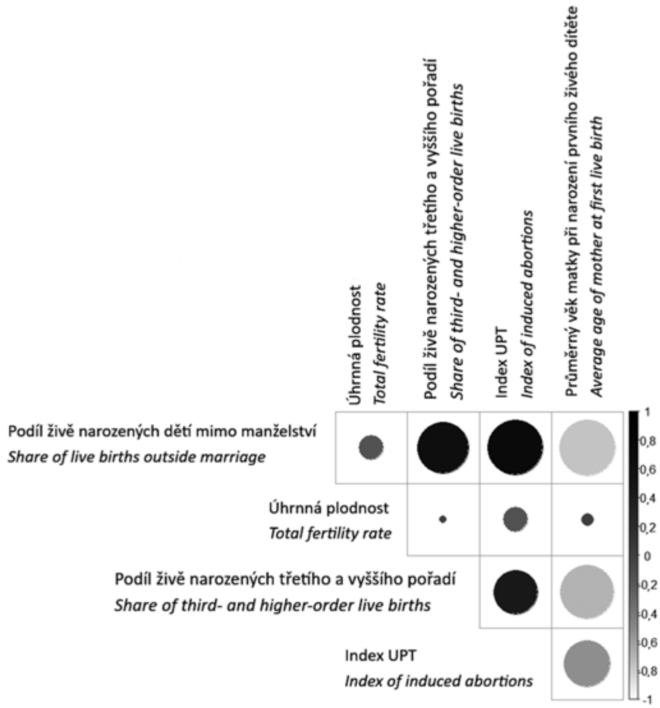
Zdroj: Vlastní zpracování na základě dat ČSÚ.
Source: *Authors' calculation based on CZSO data.*

Příloha 2: Korelační matice sledující závislost mezi proměnnými v roce 2015
Correlation matrix monitoring the dependence between variables in 2015



Zdroj: Vlastní zpracování na základě dat ČSÚ.
Source: *Authors' calculation based on CZSO data.*

Příloha 3: Korelační matice sledující závislost mezi proměnnými v roce 2019
Correlation matrix monitoring the dependence between variables in 2019



Zdroj: Vlastní zpracování na základě dat ČSÚ.
Source: *Authors' calculation based on CZSO data.*

SUMMARY

Our analysis of reproductive behaviour across Czech districts in 2015, 2019, and 2023 shows persistent and substantial spatial differences. Descriptive statistics confirm variation in both the levels and variability of key indicators, while also revealing shifts in their distribution and identifying districts that consistently diverge from national patterns. Correlation analysis indicates stable relationships between variables, reflecting the distinct regional reproductive strategies that remain visible across the entire period.

Cluster analysis divided the districts into five groups with clearly differentiated profiles. One cluster is characterised by a low mean age at first birth, a high share of nonmarital births, and a higher index of induced abortions for nonmedical reasons. Another cluster shows the opposite pattern, with older firsttime mothers and low values for the remaining indicators. A third cluster groups districts with average values, while a fourth is marked by share of third- and higher-order live births and a fifth by a distinctively high index of induced abortions for nonmedical reasons. These patterns demonstrate that reproductive behaviour in Czech districts can be differentiated into several identifiable types that are shaped by multiple demographic characteristics rather than geography alone.

The study achieved its objectives of describing regional differentiation, comparing developments across the three years, and identifying five clusters that capture how reproductive patterns are grouped across Czech districts. The findings provide an uptodate and comprehensive view of the spatial structure of reproductive behaviour and offer a solid basis for further socio-economic and cultural interpretation, especially in relation to longterm demographic change.

PODKLADY

Redakce přijímá rukopisy v elektronické podobě.

ROZSAH PŘÍSPĚVKU:

Doporučený maximální rozsah textové části článku je 50 000 znaků včetně mezer. Příspěvky do oddílů: Přehledy by neměly přesahovat 38 000 znaků vč. mezer, recenze a zprávy 10 000 znaků vč. mezer. Je třeba, aby studie obsahovala anglický abstrakt do 5 řádků (Ř), klíčová slova v angličtině, anglické resumé do 20 Ř, abecední seznam citované literatury a CV – stručnou informaci o autorovi a jeho odborném zaměření (do 5 Ř). Součástí článku je abecední seznam citované literatury.

Rukopis je třeba zaslat v textovém editoru Word, zdrojová data pro tabulky a grafy v programu Excel, obrázky a mapy ve formátu *.tif, *.jpg, *.eps. Tabulky, grafy a obrázky je třeba zařadit do textu. Názvy i těla tabulek, grafů a obrázků musí být dvojjazyčné (česko-anglické).

Recenzní řízení je oboustranně anonymní. Rozhodnutí o publikování rukopisu, resp. závěru redakční rady, je autorovi sděleno do 14 dnů po zasedání redakční rady.

Redakce provádí jazykovou úpravu textu. Anglický text je revidován rodilým mluvčím na náklady redakce.

ZÁSADY PRO OPTIMÁLNÍ PODOBU PODKLADŮ

A. TEXTY (v textovém editoru MS Word)

1. V nastavení odstavce používejte pouze zarovnání VLEVO (na levou zarážku).
2. Vyznačování v odstavci (kurzívou, tučně) a používání indexů bude do sazby korektně přeneseno.
3. Nepoužívejte (v nastavení vypněte) funkci, která nuceně přesunuje do další řádky jednohláskové předložky a spojky (a, s, z, v, k apod.), jež by jinak vyšly na konec řádku. Textový editor vsune do textu programové informace o tomto tzv. nuceném dělení, které nelze jinak než pracně odstranit.

B. GRAFY, OBRAZOVÉ SOUBORY

1. Pro zpracování grafů je kromě požadovaného typu (sloupcový, spojnicový, bodový apod.) nutné připojit zdrojová data v programu Excel.
2. Všechny obrazové soubory – např. mapy, fotografie ukládejte mimo textový soubor samostatně ve formátech *.tif, *.jpg, *.eps s odkazem v textu (graf 1, schéma 1 apod.).
3. Pro další technologické zpracování je důležité, aby bitmapové soubory měly ve velikosti 1:1 rozlišení 300 dpi.

C. PRAVIDLA CITACÍ A POPISKY

Pokud má dokument přiděleno doi, musí být v citaci uvedeno.
Příklady základních druhů citací:

Monografie

Roubíček, V. 1997. *Úvod do demografie.* Praha: Codex Bohemia. (U publikace s více než třemi autory se uvádí jen příjmení prvního autora, za ním následuje zkratka a kol., u zahraničních publikací et al.)

Hantrais, L. (ed.). 2000. *Gendered Policies in Europe. Reconciling Employment and Family Life.* London: Macmillan Press.

Potraty. 2005. Praha: Ústav zdravotnických informací a statistiky.

Články v časopisech

Dudová, R. – Vohlídalová, M. 2018. Muži a ženy pečující o seniory v rodině. Sociologický časopis, 54(2), s. 219–252. https://doi.org/10.13060/00380288.2018.54.2.400.

Články ve sbornících

Daly, M. 2004. Rodinná politika v evropských zemích. In *Perspektivy rodinné politiky v ČR,* s. 62–71. Praha: MPSV ČR.

Elektronické dokumenty

Je třeba uvést:

1. Specifikaci média (on-line, databáze, datový soubor)
2. Datum stažení (cit. 29. 10. 2005)
3. Webovou adresu (dostupné z: http://www.czso.cz)

Přednášky z konferencí

Maur, E. *Problémy studia migrací v českých zemích v raném novověku.* Příspěvek přednesený na konferenci Dějiny migrací v českých zemích v novověku. Praha, 14. 10. 2005.

Seznam literatury a odkazy

Jednotlivé položky jsou řazeny podle abecedy, více prací od téhož autora je řazeno sestupně od nejstarší k nejnovější. Pokud má autor v seznamu v jednom roce více pložek, rozlišují se přidáním písmen a, b, c... za rok vydání.

Příklad:

Syrovátka, A. 1962a. Úrazy v domácnosti. *Česká pediatrie,* 17, s. 750–753.

Syrovátka, A. 1962b. Úmrtnost dětí v českých zemích na dopravní úrazy. *Časopis lékařů českých,* 101, s. 1513–1517.

Odkazy v textu na seznam literatury

(Srb, 2004*)*; *(Srb,* 2004: 36–37*)*; *(Syrovátka a kol.,* 1984*)*. (*Dudová – Vohlídalová*, 2018)

Popisky tabulek a grafů (dodat v češtině a angličtině)

Tab. 1: Pohyb obyvatelstva, 1990–2010; Population and vital statistics, 1990–2010

Graf 1: Relativní věková struktura cizinců a obyvatelstva ČR celkem, 31. 12. 2009; Relative age distribution of foreigners and total population of CR, 31 Dec 2009

D. DOPORUČENÁ STRUKTURA ČASOPISU

Viz: https://csu.gov.cz/pokyny_pro_autory.

ČESKÝ
STATISTICKÝ
ÚŘAD

Demografie

revue pro výzkum
populačního vývoje

csu.gov.cz

Demografie, revue pro výzkum populačního vývoje
Demografie, Review for Population Research

Vydává Český statistický úřad
Published by the Czech Statistical Office

Adresa redakce: Na padesátém 81, 100 82 Praha 10, Česká Republika
Telefon: +420 274 054 240
E-mail: redakce@csu.gov.cz
Web: https://csu.gov.cz

Časopis je v plném znění uveřejněn (od roku 2004) na internetu na adrese:
https://csu.gov.cz/demografie

Objednávky vyřizuje: Informační servis, Český statistický úřad, Na padesátém 81, 100 82 Praha 10-Strašnice, Česká republika, e-mail: objednavky@csu.gov.cz

Grafická úprava: Český statistický úřad
Grafický návrh: Ondřej Pazdera
Tisk: Český statistický úřad
Cena jednoho výtisku: 58,- Kč
Roční předplatné 4 x 58,- Kč + poštovné

Indexové číslo 46 465, ISSN 0011-8265 (Print), ISSN 1805-2991 (Online),
Reg. Zn. MK ČR E 4781

Číslo 1/2026, ročník 68
Toto číslo vyšlo v březnu 2026