Identification of Unregistered Emigration in the Norwegian Population Register

Stian Nergård, Linn Krokedal

Statistics Norway, Norway

Abstract

A precise estimate of the target population is inherently important in population statistics. However, factors such as increased immigration, and few incentives for deregistration after emigration mean that population registers may not always accurately reflect the target population. This study aims to identify unregistered emigration using "signs of life". That is, detecting historical inactivity of individuals who have emigrated, but are still listed as residents in the population register. Unregistered emigration contributes to over-coverage, as the number of actual emigrations exceeds the number of registered emigrants. This estimation error affects not only size and composition of the population, but also impacts demographic indicators, such as mortality and fertility rates. Statistics on households and families may also become skewed due to these discrepancies. There is still no consensus on how to identify or deal with unregistered emigration. Addressing this, we first provide a comparison of methods adapted from the literature for estimating the number of unregistered emigrations. The zero income approach provides a method with minimal computational and data quality requirements, which serves as a foundation for the other estimation techniques. The household income approach builds upon this by correcting for household income factors. Finally, the register trace approach provides the most comprehensive and detailed picture of unregistered emigration. Our estimates suggest that unregistered emigrants account for approximately 0.44 percent of the adult population in Norway. Second, we analyse the demographic characteristics of the non-deregistration group. We find that the problem of unregistered emigration is not equally distributed across the population, indicating that some subgroups are more prone to discrepancies than the rest of the population. Among immigrants, the over-coverage due to unregistered emigration is substantially higher, accounting for 2.29 percent of the population.

Keywords: Over-coverage, Population register, Emigration, Norway

1.   Introduction

This paper describes Statistics Norway's (SSB) methodology for estimating the prevalence of unregistered emigration in the Norwegian National Population Register (NPR). It expands upon a note published by SSB in Norwegian (Krokedal et al., 2024). We define unregistered emigrants as individuals who have emigrated from Norway and should be registered as emigrated, but who remain listed as residents because the emigration was not registered.

In Norway, any migration event or change of address intended to last more than six months must be registered with the NPR by law (Skatteetaten, 2024). However, the registration of emigration is less reliable than for other events. The low compliance is likely a result of there being few incentives to register emigration and/or a lack of knowledge about the registration

process. Under-reporting of emigration leads to over-coverage in the population, which impacts population statistics in several ways. The estimated population will be higher than the target population of people residing in Norway, all else equal. The over-coverage may also lead to inaccuracies in demographic rates, such as birth- and mortality rates (Monti et al., 2020). Moreover, with a selection effect in the group of unregistered emigrants, these rates will be more biased in some subsets of the population. Population forecasts, and by extension "academic research, marked analysis, public policy, and planning" (Jia et al., 2023), could also be impacted by over-coverage.

There is ongoing work in the NPR on using control measures to identify unregistered emigration, a process that can lead to so-called administrative emigration. This refers to the NPR "emigrating" people who are still listed as residents, but are believed to no longer be living in Norway. In 2023, of the near 34,000 registered emigration events, around 12,000 were due to administrative emigration, a share of 35 percent. This number was 47 and 44 and percent in 2021 and 2022, respectively. Because of the personal consequences of being mistakenly marked as emigrated, the NPR have strict identification criteria for administrative emigration. Additionally, there are numerous legislative barriers associated with this process, meaning the control measures may change over time. Statistical emigration, which is the process of flagging unregistered emigrants and excluding these individuals from aggregate statistics, does not have the same personal consequences for those identified, nor the same legislative obstacles. As a result, it differs from administrative emigration, and can provide a consistent way to identify unregistered emigration which can be readily integrated in population statistics.

We first provide an overview of the relevant literature, including previous work by SSB. Then, in Chapter 3, we describe the methods and data we have used to estimate unregistered emigration in Norway. Next, in Chapter 4 we compare the estimation methods, evaluate the procedure, and discuss the results. Finally, in Chapter 5 we provide a conclusion.

2. Literature Review

Over-coverage has become more relevant in recent years, as more countries move to register-based systems in collection of population data and census reporting (Poulain & Herm, 2013; Skinner, 2018). Although the issue of unregistered emigration as a source of over-coverage is well documented, there is no consensus among national statistical offices on how to identify and correct this issue.

Between 1995 and 2010 Statistics Sweden (SCB) produced eight different studies of over-coverage in the Swedish population register. From 2004 yearly estimates of unregistered emigration have been provided using a "register trace approach" (Statistics Sweden, 2015).

Ludvigsson et al. (2016) estimate over-coverage to be between 0.25-0.50 percent. Since then, several studies have been carried out using Swedish register data to compare methods of estimation and estimate the effect of over-coverage on demographic rates (Monti et al., 2020), as well as refining the method for identifying unregistered emigration to probabilistic models using multiple systems estimation (MSE) (Mussino et al., 2023). In Statistics Finland, individuals identified as unregistered emigrants using a "signs of life" approach are removed from the statistical population (Statistics Finland, 2024).

In Norway, there have been several studies identifying the issue of over-coverage, by for instance exploring weak connections to the labour market (i.e., Horgen, 2014; Bø & Vigran, 2015) or using unregistered emigration as an explanation of the differences in mortality rates between immigrants and the host population (Syse et al., 2016). Krokedal et al. (2024) provide a generalised approach to estimating unregistered emigration and find around 0.4 percent over-coverage from unregistered emigration in the Norwegian population register using a register trace approach similar to that used in SCB.

3. Method and Data

In this paper, we utilise cross-sectional data on residents in Norway, aged 15 or older, at the start of the calendar year. On January 1st, 2023, this group comprised around 4,570,000 people. We also use annual data from administrative registers, including data on income, residential addresses, demographic events, labour market activity, welfare payments, and education to track activity. Activity in any of these data sources indicates that a person is residing in Norway. Conversely, a lack of activity suggests that the person may have emigrated without registering the event in the NPR.

3.1 Basic Theoretical Framework

Let P_t be the population from which we search for inactivity. In this case, P_t is the set of the population of Norway aged 15 or older, at the beginning of the year t, where n is the size of the population in question.

$$P_t = \{p_1, p_2, \dots, p_n\}$$

Let R_{t-1} be the set of administrative data containing information on any individual's activity in the previous year $t-1$. We define the indicator function, $I(p_i, R_{t-1,j})$, for any element $p_i \in P_t$, with respect to each of the data sources $R_{t-1,j} \in R_{t-1}$ as follows:

$$I(p_i, R_{t-1,j}) = \begin{cases} 1, & \text{if } p_i \notin R_{t-1,j} \\ 0, & \text{if } p_i \in R_{t-1,j} \end{cases}$$

This function returns the value one if a person had no activity in a specified data source over the past year, and zero if there was activity. The sum of $I(p_i, R_{t-1,j})$, over all individuals in P_t, gives an estimate of the number of unregistered emigrants. The method can be extended to incorporate multiple criteria by taking the sum of the products of indicator functions for each data source, effectively evaluating whether p_i is not in any of the datasets $R_{t-1,j}$.

3.2 Zero Income Approach

One way to estimate the number of unregistered emigrants is to identify residents with no personal income, using a so-called zero income approach (i.e., Aradhya et al., 2017; Monti et al., 2020). The income variables considered are labour income, capital income, taxable transfers, and non-taxable transfers. Any person registered as a resident in Norway at the start of the year, who did not receive any personal income the previous year, is flagged as a potential unregistered emigrant. However, there are some limitations to this model. For example, people could receive financial support from family or friends, earn income from undeclared work, or receive earnings from abroad. Moreover, the estimation technique does not account for children, who generally do not have any personal income.

Define PI_{t-1} as the set of all p_i with registered personal income strictly greater than zero. Then, the estimated number of unregistered emigrants, M, using the zero income approach is:

$$\text{M} = \sum_{i=1}^{n} I(p_i, PI_{t-1})$$

3.3 Zero Income Approach Corrected for Household Income

The second model attempts to correct some of the limitations of the zero income approach by also considering household income instead of just personal income. A household is defined as individuals who are permanently residents of a common dwelling and have a shared economy (Statistics Norway, 2023). Household income is the sum of the personal incomes of the members of the household. A person is only flagged as a potential unregistered emigrant if they live in a household with a total income lower than what can be considered possible to live off in Norway. This threshold is based on the base-number (G) for calculating payments in the national insurance scheme. In 2023, this was 118,620 Norwegian kroner (NAV, 2024). The threshold is adjusted for number of people in the household.

This estimation procedure depends on accurate information about households. Unregistered emigration may itself be a complicating factor. If someone emigrates without

registering with the NPR, another family may move into the same dwelling, which means the new family's total income may be sufficiently high to cover the emigrated family with no income. We attempt to correct for this by identifying what we call deviating households as potential unregistered emigrants. A deviating household is defined as individuals in multiple-family households who have a moving-in date more than a year earlier than the most recent moving-in date in the household, removing institutions and other addresses where we expect multiple families with varying moving-in dates.

Define HI_{t-1} as the set of all p_i with either a low household income or who are in a deviating household. The estimated number of unregistered emigrants, M, accounting for household income and deviating households is calculated as follows:

$$M = \sum_{i=1}^{n} \left(I(p_i, PI_{t-1}) \cdot I(p_i, HI_{t-1}) \right)$$

The term $I(p_i, PI_t)$ checks for zero personal income, while $I(p_i, HI_t)$ checks whether p_i has a substantially low household income *or* lives in a deviating household. Multiplying the terms means that the estimate includes people with inactivity in both data sources.

3.4 Register Trace Approach

The third model combines income data with activity from different linked registers, including the NPR and the System for Personal Data (SFP). From the NPR we collect data on demographic events such as giving birth (including the father), changes of address, and changes in marital status. The SFP comprises several administrative registers and contains information on labour market activity, education, and benefit payments. The register trace approach assumes that unregistered emigrants, having left the country, do not leave any 'traces' in these administrative registers. The income-based approaches only pick up on some of these traces and therefore may misidentify some people as unregistered emigrants. In this approach, unregistered emigrants are defined as those with zero personal income, low household income (including those in deviating households), and with no activity in the administrative registers. Formally, we can represent this as:

$$M = \sum_{i=1}^{n} \left(I(p_i, PI_{t-1}) \cdot I(p_i, HI_{t-1}) \cdot I(p_i, NPR_{t-1}) \cdot I(p_i, SFP_{t-1}) \right)$$

4.   Results

4.1 Comparison of Estimation Approaches

The three approaches – zero income, zero income corrected for household income, and register trace – provide different estimates for the prevalence of unregistered emigration. Consequently, the choice of method impacts both who, and how many, are statistically emigrated. In the following chapter, we will compare and explain the differences between the outcomes in the various approaches.

Table 3: Estimated number of unregistered emigrants by year and model.

	2021	2022	2023
Zero Income Approach	90,028	98 108	99,199
Zero Income Corrected for Household Income	32,182	38 224	44,068
Register Trace Approach	20,929	23 570	20,175

On January 1st, 2023, nearly 100,000 registered residents, aged 15 or older, had no registered personal income in 2022. A significant proportion of this group were young people; nearly half of them (47,937) were younger than 18. Most of these teenagers live with their parents and still go to school, making it unlikely that they are unregistered emigrants. While the zero income approach is simple and requires few data sources, only using this approach is likely too simplistic in identifying unregistered emigrants. It is still worth noting the number of people who don't have any personal income, as zero income is a criterion for being assessed as an unregistered emigrant in the methods used in this paper.

The estimated number of unregistered emigrants drops considerably when we include household income. In 2023, the estimate declined from 99,000 to around 44,000, equalling a drop of 56 percent. The large correction of the zero income approach can be attributed to identifying young individuals who are not emigrated. For instance, the estimated unregistered emigrants under the age of 18 in 2023 dropped from around 48,000 to around 7,000. Thus, there is good reason to believe that including household income gives us a better picture of who may be an unregistered emigrant. However, it is important to note that if certain household members emigrate without registering this, but the rest of the household remains, the inclusion of household income can be a source of error. Nevertheless, this is a less frequent source of error than not including household income.

The inclusion of data from the NPR and SFP reduces the estimated number of unregistered emigrants in 2023 by approximately 24,000 people, corresponding to a

percentage change of 54. This is an indication that the administrative registers are effectively identifying individuals with no personal income, a low household income or who are living in a deviating household, who nevertheless reside in Norway.

4.2 Discussion and Evaluation

Having zero registered personal income is the first criteria for identification as an unregistered emigrant in our estimates. However, this may lead to an underestimation of the number of unregistered emigrants. Looking at the data we find that many individuals have no other traces of activity in any of the administrative registers yet have a low non-zero personal income. If we increase the income threshold to G, the number of estimated unregistered emigrants in 2023 would have been 22,823. A likely explanation for this is that the broad income definition means that individuals can emigrate from Norway and still receive some personal income, such as holiday pay, determined as a percentage of income last year (The Norwegian Labour Inspection Authority, n.d.).

Estimating unregistered emigrants for past years allows us to "evaluate" our method. One test we can do is to check our estimates from 2022 against NPR's work on administrative emigration. We find that around 7,500 of the estimated 23,570 unregistered emigrants at the start of 2022 were administratively emigrated in 2022. Approximately 12,000 were still suspects of unregistered emigration in 2023, which is natural since the process of administrative emigration takes time. The remaining 4,000 those who were identified as unregistered emigrants in 2022 had some registered activity in the following year and were not suspects of unregistered emigration in 2023. Supposing that those identified as unregistered emigrants in the next year will be administratively emigrated in the future, then the precision of the estimation method is relatively high. However, of the 14,000 administratively emigrated individuals in 2022, we are only able to identify around 7,500. Thus, the work of administrative emigration by the Tax Authority continues to be important for the quality of the population register and by extension population statistics in Norway.

4.3 Descriptive Statistics on the Estimated Enregistered Emigrants

In total, there are 20,175 people who, according to the estimate from the register trace approach of 2023, are assumed to have emigrated without being officially registered as such. They made up 0.44 percent of all residents in Norway aged 15 and older, as of January 1st, 2023. Among those who are presumed to have emigrated, most are immigrants. They make up 85 percent of the group, which corresponds to 17,108 people. Although most are labour immigrants (39 percent), there is a significant number of immigrants with family-reunification

(28 percent) and education (15 percent) as their reason for immigration. The largest group by country-background of unregistered emigrants are born in Poland, they make up a group of 2,899. People born in Norway make up the second largest group with 2,471 people, followed by people born in Lithuania, Somalia, and Germany. Around 30 percent had a registered period of residence of less than 4 years as of January 1st, 2023, around 21 percent had a period of residence between 5 and 9 years, while 49 percent had a period of residence longer than 10 years. However, they may have emigrated some time before this, so that the actual length of stay in Norway is probably somewhat shorter. Men make up 57 percent of presumed emigrants. Among migrant workers, men make up 78 percent, which explains much of the gender distribution. On the other hand, women are overrepresented in groups where reason for immigration is family-reunification and education. There are only 2,471 people among the presumed emigrants who have not immigrated, of these 1,491 are men (60 percent).

115 of the presumed emigrants are older than a hundred years. This corresponds to 0.6 percent of the presumed emigrants and constitutes a far higher proportion, of people a hundred years or older, than for the general population. These people represent an interesting issue; if someone emigrating does not register emigration, information about their death will also not be received. Thus, these individuals can in theory live in the registers forever. People with an improbable age and no activity in the registers, may therefore have died in another country. These people are sometimes referred to as "administrative survivors".

5.   Conclusion

Unregistered emigration is a problem that challenges the accuracy of register-based population statistics. It is nevertheless important to specify that the NPR maintains a high level of data quality internationally. For persons aged 15 or older, we find an over-coverage resulting from unregistered emigration which corresponds to 0.44 percent of the population in Norway. Most of these were young men from Eastern Europe who came to Norway as migrant workers.

Unregistered emigration may mean, for statistical agencies, that the register population does not match the actual population. It is therefore advantageous to find good solutions through regulations and/or controls so that the register population in the future will be as accurate as possible. In this way, demographic estimates, and projections, as well as all other statistics that use the population figure, will be of a higher quality. In our paper, we have considered different models that can estimate unregistered emigration, with varying requirements for data quality and availability. Based on the models we have looked at, we find that the register model gives the most accurate representation of the prevalence of unregistered emigration in Norway, given the data sources we have.

References

Aradhya, S., Scott, K., & Smith, C. D. (2017). Repeat immigration: A previously unobserved source of heterogeneity? *Scandinavian Journal of Public Health, 45(17_suppl)*, pp. 25-29. doi:10.1177/1403494817702334

Bø, T. P., & Vigran, Å. (2015). *Ungdom som verken er i arbeid eller utdanning.* Oslo/Kongsvinger: Statistics Norway.

Horgen, E. H. (2014). *Innvandrere med svak tilknytning til arbeidsmarkedet - hvem er de?* Oslo/Kongsvinger: Statistics Norway.

Jia, Z., Leknes, S., & Løkken, S. A. (2023). Moving beyond expectations: From cohort-component to microsimulation projections. *Discussion Papers No. 999, Statistics Norway*.

Krokedal, L., Nergård, S., & Kvalø, E. (2024, januar 15). *Uregistrerte utvandringer fra Norge.* Oslo/Kongsvinger: Statistics Norway. Retrieved from ssb.no.

Ludvigsson, J. F., Almqvist, C., Bonamy, A.-K. E., Ljung, R., Michaëlsson, K., Neovius, M., . . . Ye, W. (2016). Registers of the Swedish total population and their use in medical research. *European Journal of Epidemiology, 31(2)*, pp. 125-136. doi:10.1007/s10654-016-0117-y

Monti, A., Drefahl, S., Mussino, E., & Härkönen, J. (2020). Over-coverage in population registers leads to bias in demographic estimates. *A Journal of Demography, 74(3)*, pp. 451-469. doi:10.1080/00324728.2019.1683219

Mussino, E., Santos, B., Monti, A., Matechou, E., & Drefahl, S. (2023). Multiple systems estimation for studying over-coverage and its heterogeneity in population registers. *Quality & Quantity*. doi:10.1007/s11135-023-01757-x

NAV. (2024). *Grunnbeløpet i folketrygden*. Retrieved from nav.no: https://www.nav.no/grunnbelopet

Poulain, M., & Herm, A. (2013). Central Population Registers as a Source of Demographic Statistics in Europe. *Population, 68(2)*, pp. 183-212. doi:10.3917/popu.1302.0215

Skatteetaten. (2024). *Moving from Norway*. Retrieved from skatteetaten.no: https://www.skatteetaten.no/en/person/national-registry/moving/moving-from-norway/

Skinner, C. J. (2018). Issues and Challenges in Census Taking. *Annual Review of Statistics and Its Application, 5*, pp. 49-63. doi:10.1146/annurev-statistics-041715-033713

Statistics Finland. (2024). *Preliminary population statistics: documentation of statistics*. Retrieved from stat.fi: https://stat.fi/en/statistics/documentation/vamuu

Statistics Norway. (2023). *Families and households*. Retrieved from ssb.no: https://www.ssb.no/en/befolkning/barn-familier-og-husholdninger/statistikk/familier-og-husholdninger

Statistics Sweden. (2015). *Övertäckningen i Registret över totalbefolkningen.* Statistics Sweden.

Syse, A., Strand, B. H., Naess, O., Steingrímsdóttir, Ó. A., & Kumar, B. N. (2016). Differences in all-cause mortality: A comparison between immigrants and the host population in Norway 1990-2012. *Demographic Research, 34*, pp. 615-656. doi:10.4054/DemRes.2016.34.22

The Norwegian Labour Inspection Authority. (n.d.). *Holiday pay*. Retrieved from arbeidstilsynet.no: https://www.arbeidstilsynet.no/en/working-conditions/holiday/holiday-pay/