BI Norwegian Business School - campus Oslo

GRA 19703

Master Thesis

Thesis Master of Science

An empirical analysis of housing allowance recipients 2010-2020 and a forecast of the near future

Navn: Henrik Jenssen, Jon Peder Bakke Unhjem

Start: 15.01.2020 09.00

Finish: 01.09.2020 12.00

Abstract

This study seeks to identify the attributes of recipients of housing allowance in Norway and formulate an econometric model capable of predicting the inflow of new recipients of housing allowance in the near future. As the Norwegian State Housing Bank receives funding via the state budget, such a model will help to ensure that the bank receives proper funding.

The study finds that the number of applicants, recipients and new recipients has decreased significantly over the past ten years, despite that the number of people considered poor in Norway has increased. Moreover, the anticipated effect of income and housing expenses are minimalised through the politically decided income and approved housing expenses limits. Furthermore, we find the inflow of new recipients to be a function of previous inflow, average housing expenses, age, regulations, employment, and unemployment. The chosen model to forecast the inflow is Vector Autoregression (VAR) model. Moreover, through an Impulse Response Function (IRF), we find that Regulation and Employment are the two variables that has the greatest effect on the inflow of new recipients. The accuracy of the model is tested by comparing the VAR forecast to a forecast with linear regression and actual values. Moreover, it is evaluated using mean error (ME), mean percentage error (MPE) root mean squared errors (RMSE), mean absolute errors (MAE), mean absolute percentage errors (MAPE).

The study concludes that given the available data, the VAR model is able to produce satisfactory results, although the precision and the usage of external data can be better.

Acknowledgements

This study is conducted as part of the Master of Science program in Business Analytics at BI Norwegian Business School.

We would like to thank our supervisor Jon H. Fiva for guiding us through the process. Moreover, we would like thank Hero Ezat, Andreas Fjelltoft, and the rest of the Norwegian State Housing Bank for input, data, and advice.

Finally, we would like to thank our family and friends for their continuous and valuable support.

Table of content

Table of Figures

Table of tables

1. Introduction

According the World Bank Group (2020), Norway is the tenth richest country in the world when measuring GDP per capita. However, poverty amongst the inhabitants in Norway has increased in recent years (SSB, 2019). To cope with poverty, the Norwegian welfare state has numerous schemes to provide its inhabitants with economic help and social wellbeing.

In this report we will examine one of these schemes, housing allowance. The main purpose of housing allowance is to help families and individuals obtain and/or maintain a satisfactory living situation. Housing allowance is distributed by The Norwegian State Housing Bank (NSHB).

Our research will build on the analysis conducted by Fjelltoft & Ezat (2019). They created a picture of who the recipients were in February 2019 and how they had developed since the reform in 2009. However, they did this by looking at one month per year. We will divide our report into two main parts. The first part will be a description of who the recipients are and how the development of the recipients has been in the period 2010-2020. The biggest difference from Fjelltoft & Ezat's (2019) work will be that we will utilize every month from January 2010 – March 2020 to explain who recipients are and to show how they have developed. In the second part of our thesis we will, based on a Vector Autoregression (VAR) model on an aggregated multivariate time series and an Impulse Response Function (IRF), look at how extreme changes (shocks) in both external and internal variables will affect the inflow of new recipients. The same model will also be used to predict the inflow of new recipients from 1st April 2020 – 31st December 2021. Through these analyses, we aim to answer the following research questions:

1. Who are the recipients of housing allowance and how have they developed from 2010-2020?
2. How does the inflow of new recipients get affected by shocks on internal and external variables?
3. How will the inflow of new recipients of housing allowance be in the near future?

The intention of this research is that it will contribute to a better understanding of the recipients of housing allowance and the nuances that separates them. The report will provide a better knowledge base for further work on developing an accurate forecast model for the inflow of total- and new recipients.

We found that the total number of applications, the total number of recipients and total number of new recipients all have downward trends. When comparing new recipients to recipients already in the system, new recipients are younger, have higher housing expenses and lower income. In the second part of the analysis, we found that regulation and employment had the biggest effect on the inflow of new recipients. Furthermore, the downward trend we previously found in the descriptive part of the paper will continue until 31st of December 2021.

The paper is divided into 7 main chapters, including this introduction. In chapter 2 we will expand on the institutional setting of which housing allowance operates. In chapter 3 we will expand on the main body of relevant research and literature related to our paper. In chapter 4 we will introduce our data. Chapter 5 contains the methodology used to create and test our model. In chapter 6 we will display the results of our data analysis. Lastly, in chapter 7 you will find our conclusion and final thoughts of the experiment.

2. Institutional setting

The practice of Housing allowance and the importance of it cannot be understood on its own. To fully value and comprehend it, it must be viewed upon in the context of the Norwegian Society and values as a whole.

2.1 The Welfare State

Most modern countries practice some elements of what is considered a "welfare state". In the broad sense welfare state means a type of governing in which the national government protect and promote economic and social well-being of its citizens. The core values of the welfare state are equality of opportunity, equitable distribution of wealth, and public responsibility for those unable to avail themselves of the minimal provisions of a decent life. Social security, welfare payments, and free public education are all examples of a welfare state (Kenton, 2019). These schemes are usually paid for through taxation of individuals as well as companies (Christensen & Berg, 2019).

2.2 The Norwegian Welfare State

The Norwegian welfare state is based upon the general values described in the previous chapter, is wide-ranging and has roots dating back to the 18th and 19th centuries when the government relieved individual families and the church from this responsibility. The goal was to help those who lived in poverty and take action against social distress. However, the payments were minimal and very stigmatizing (Christensen & Berg, 2019). The development of the welfare state with social policies as we know it today were introduced in the beginning of the 20th century and only gained momentum after World War II ended. Child benefit, sickness benefit, unemployment benefit and general old-age benefit were introduced in quick succession and collected in the National Insurance Scheme in the mid-1960s.

Today, the services can broadly be split into two categories, by universal and individual means. The universal rights are mostly limited to services such as healthcare and education. This entails that regardless of your income you qualify for free, or a small deductible, healthcare, and education. Public transfer-schemes such as housing allowance and several other benefits are scaled by income.

A quintessential principle for the Norwegian government is that the public benefits must maintain a high standard. This is to ensure that people with higher income does not opt to buy private services. The fact that most people, regardless of income, use public services is a prerequisite for support for the welfare state.

The Norwegian welfare state is largely financed through general taxation, i.e. taxes on income, consumption and wealth that are not directly linked to welfare benefits. As those with the highest income pay the most in taxes, while public benefits are relatively evenly distributed across the population, the Norwegian welfare state has a significant redistributive effect. A comprehensive welfare state is one of the most important reasons why income inequalities in Norway are far smaller than in many other countries (Christensen & Berg, 2019).

2.3 Housing Policy: Adequate and secure housing for all

The primary vision for the Norwegian housing policy is adequate and secure housing for all. To achieve this vision, The Norwegian State Housing Bank, the Norwegian Building Authority and the Rent Disputes Tribunal all work closely with the department for Housing and Building, which in turn is part of the Ministry of Local Government and Modernisation. Amongst the many responsibilities, the Ministry is responsible for implementing the government's housing and building policy. Other responsibilities are the Planning and Building Act, local government finances and local administration, ICT Policy and Public Sector Reform, rural and regional policy, the conduct of elections, government employer policy, Sami and minority affairs and national mapping and geodata policy.

The main goals of the housing and building policy are; Houses/ homes for everyone, in good living environments, security of tenure in owned and rental homes, housing conditions that promote prosperity and participation in society, well-designed, secure, energy-efficient and healthy buildings, better and more efficient construction processes.

To achieve the aforementioned goals there are five main instruments. Legislation, information, housing allowances, grants to help people to obtain their own home, and loans and grants for building and upgrading housing units.

Through legislation the department provides guidelines in order to facilitate a functioning housing and building market. This means that it is the department's responsibility to ensure a sound and effective legal framework, an efficient and fair system for dealing with building matters, slowing down the inflation of building costs, providing a high level of expertise and an efficient framework and inspection system for the building industry.

Furthermore, through housing allowance and grants the department can provide homes for those who initially are unable to access the housing market, i.e. low-income households such as refugees, persons with disabilities and persons who have fallen out of employment. This is seen as one of the most important tools the department control in order to combat poverty and homelessness.

Moreover, by providing loans and grants for building and upgrading housing units the department are able to promote sustainable quality, security and high aesthetic standards in the built environment.

Lastly, the department provides information and promotes awareness and knowledge of good building practices and sound urban settlement development. This to reduce building errors and building faults and damages. Further, the department promotes awareness and knowledge of universal design among consumers, local government authorities and key actors involved in the building process. This to achieve that a larger number of homes, buildings and outdoor spaces are based on universal design principles.

2.4 The Agents

There are three principle agents which cooperates in providing housing for the Norwegian population: The National Government, municipal authorities and private firms and organizations.

The Norwegian State Housing Bank (NSHB) is the implementing agency and provides loans, grants, and guidance as well as initiates new development and research. Whereas the government and the parliament are responsible for overall housing policy goals, as well as the financial- and legislative framework. As the main housing policy is adequate and secure housing for all, the municipalities are responsible for ensuring that the disadvantaged have access to adequate and good housing. The municipalities are also responsible for enabling building and rehabilitation of both public and private property (Norwegian State Housing Bank, ND).

2.5 About the Norwegian State Housing Bank (NSHB)

NSHB was established by the Parliament in 1946. They use a set of financial tools and aids to facilitate the achievement of the government's housing policy goals. The most important financial tools the NSHB has available are basic loans, start-up loans, housing grants and housing allowances. Moreover, The NSHB administers the government compensation scheme for renovation of schools and churches, provides loans to day-care centres and subsidises student housing. The NSHB we encounter today is also a resource centre for housing policy matters and takes an active role in providing information and guidelines and promoting knowledge development. (Norwegian State Housing Bank, ND).

2.5.1 History

As described in the introduction of this chapter, the NSHB was founded in 1946, a year after WWII ended. Much of northern Norway was in ruins after Germany's scorched earth strategy. However, even before WWII, the housing was precarious in the districts and towns at the turn of the previous century. The combination of this lag in housing development and WWII lead to an overwhelming lack of housing all over Norway. Furthermore, due to the war there was a significant lack of private capital and credit. Therefore, it was natural to establish a state-owned housing bank with the main purpose of "providing central and local government support for reconstruction and new building."

Since its foundation in 1946, NSHB has remained the Norwegian government’s most important tool to implement its housing policy and has played a key role in the development of the Norwegian welfare state.

The housing bank as we know it today was formed in the mid-90s. when the transition from housing-bank to welfare institution and centre for expertise begun. NSHB has moved away from the financing of new homes and started helping the disadvantaged in the housing market. Loans were issued without subsidies, while grant schemes and housing allowance were aimed at specific groups. In 2003, it was decided that the NSHB was to expand their business into helping young people and the disadvantaged with an opportunity to establish themselves in their own homes through a start-up loan. However, in 2014, The Government changed their strategy. The Start-up loan was now aimed at people with long-term difficulties, rather than first time byers who could obtain loans from private banks. This to further their goal of adequate and secure housing for all (Norwegian State Housing Bank, ND).

2.6 Housing Allowance

Housing allowance is a crucial instrument in the Norwegian housing policy. The main purpose is to help families and individuals to obtain and/or maintain a satisfactory living situation (Norwegian State Housing Bank, ND). Specifically, the state housing allowance ensures that households with low incomes receive a supplement to their main income, thereby enables them to handle their housing expenses.

The Housing allowance mechanism has two important implications. Firstly, the housing allowance can provide an incentive to ensure satisfactory housing conditions (increase housing consumption). Secondly, it can act as income protection (for households which already has satisfactory housing consumption). In this way the housing allowance will act as a safety net for households who, for various reasons, experience an unexpected decline in their income. This applies for both, short- and long term.

The insurance function of the housing allowance entails that the total benefit of this practice cannot be valued or understood on its own. To fully value and comprehend the housing allowance, it must be viewed in the context of those who receive it and those who do not, but potentially could have received it. Those who do not receive it, but do not have it as an insurance function if random events such as the corona pandemic will result in a substantial loss of income or a sudden increase in housing expenses.

2.6.1 Calculation

To determine how much a household receives in housing allowance, there is a standard formula. The process is automated, and the rates and levels of each component are politically decided. In this section, we will explain the different parts of formula, which is structured like a GAP-equation. Moreover, we will illustrate how the formula functions with an example. The formula:

Housing allowance = (approved housing expenses – deductible expenses) * Coverage percentage

i) Approved housing expenses

  a. Costs such as rent, maintenance, and mortgage-payment

  b. There is a maximum limit for approved housing expenses, this varies from municipality groups.

ii) Deductibles

  a. The deductibles are the amount of the expenses that the household is expected to pay themselves. This is dependent on the income and has 2 upper limits.

iii) Coverage percentage

  a. The coverage percentage is a fixed rate that shows how much of the difference between approved housing expenses and calculated deductible that the housing allowance covers. The coverage percentage was 73,7% in 2019.

Example: let's assume that approved housing costs are 100.000 and deductibles are 80.000 (based on income and household composition). If we fill this into the equation above, we get this:

Housing allowance = (100.000-80.000) * 0,737

In this case annually housing support amounts to NOK 14.740 per year (NOK 1.228 per month). The housing allowance thus amounts to 73.7% the difference between approved housing expenses and the deductibles.

2.9 Beneficiaries

Most of the beneficiaries of the housing allowance have disproportionately high costs in regard to their housing situation in relation to their income. We will come back to this later in the thesis. The housing allowance plays an important role in making sure that the beneficiaries can pay their housing expenses. Furthermore, it gives households with limited funds financial leeway to have a satisfactory consumption beyond housing.

3. Literature Review

In this chapter we will expand on the main body of relevant research and literature related to our paper. We will start with mapping out papers focusing on the recipients of housing allowance. Then we will present the forecasting method we applied to our multivariate time series, Vector Autoregression (VAR). Lastly, we will present literature concerning the evaluation of our method.

3.1 Recipients of Housing Allowance

The foundation for our master thesis is the study conducted by the NSHB in 2019 - (Fjelltoft & Ezat, 2019). The study is in Norwegian and called Dagens Bostøttemottakere, translated it is Current housing allowance recipients. The report aimed to identify the different groups of recipients and contribute to a better understanding of the social housing instruments. How the scheme hits in regard to purpose and seen in the context of the target group. To conduct their research, they studied data from every February in the period 2010-2019. In their conclusion they attempt, based on their findings, to look at how the scheme will develop in the years to come:

- There will continue to be fewer recipients with social security benefits that follow the National Insurance basic amount (G)
- Remaining recipients will to a greater extent have very low or no income

However, as the authors themselves mention, their analysis is in many cases based on a simplified approach, and some conclusions must be seen in the light of this.

In our study we will use an extended version of the dataset used by Fjelltoft & Ezat (2019). Our dataset will include data from January 2010 to March 2020. Some of the findings made in our study may be similar to their findings. However, this is only the case in the descriptive part of our study regarding total number of applicants. We will also focus on what characterizes the new recipients, and what separates them from the total group. Their attempts to look at how the scheme will develop in the years to come; we will use as our hypothesis' in the forecasting part of this study.

A study of whether households that qualify for the national housing benefit receives it was done by (SSB,2019). They made three estimates for eligibility based on three different criteria. To make their estimations they used a linear regression model. In the first estimate the only criteria was income. In the second estimate the criteria were housing costs. The housing costs was based on information from the SILC-survey. In the third estimate households with an income of below 50.000 were included. They found that 163,000 households qualify for housing benefits if only the income requirement is used. This is 6.9 per cent of all households. 143,000, or 6.1 per cent, qualify according to estimate 2, and 123,000 or 5.2 per cent according to estimate 3.

Although this study is interesting, it does not provide the NSHB with an estimation of future recipients nor is that its intention. This is because the study predicts how many that qualifies for housing allowance, not the ones that apply. In a "perfect world" all those who qualify for housing allowance should receive it. However, potential recipients do not apply for various reasons. Potential reasons could be lack of information, knowledge or even pride and stigma. Our study will use data provided by NSHB to first detect trends and patterns in the current and past actual recipients before we estimate future recipients based on these findings.

Moreover, the study from SSB predicts potential recipients for 2016. In 2016 the income basis was yearly, whereas after the change in 2017 the income basis became monthly. In our study we will include both numbers from before the change and after to see if there is a change in trend that will be significant in determining the future number of recipients.

3.2 VAR models

Vector Autoregression models are simple multivariate models in which each variable is explained by its own past values and the past values of all the other variables in the dataset (Holden, 1995). The extensive use of VAR models for forecasting is tribute to the work of Sims (1980).

Holden (1995) points out that there are one obvious problem concerning the general VAR model, and that is the large number of parameters that need to be estimated. He draws an example from Sims (1980) where the models have six variables and the lag length is initially eight, giving in each equation 48 coefficients excluding the constant term. On the other hand, according to Brooks and Tsolacos (2010) one of the advantages of VAR modelling is that all the variables are endogenous. This entails that we will be able to look at the effect from several variables on average inflow of new recipients. Moreover, the effect on itself, univariately. This allows us to capture more features of the data. Lastly, we can apply OLS separately on each equation.

We have not seen any studies were VAR modelling are applied in a welfare study to predict the future. The closest we have come is the aforementioned eligibility study conducted by SSB (2019) which used linear regression. However, there are several studies conducted using VAR modelling, we are confident that the principles presented in those studies are applicable to our study.

3.3 Evaluating the model

To evaluate our forecast model, we will use accuracy measurements. Accuracy measurements are usually defined on the forecast errors. For instance, mean error (ME) provide measures of bias, which is one component of accuracy (Diebold & Lopez, 1995). Furthermore, the authors states that the most common measurement of statistical accuracy in a forecasting model is the mean squared error (MSE), which in turn yield the RMSE, root mean of squared error, and that while not as popular the mean absolute error, MAE, is common.

Both the root mean of squared error (RMSE) and the mean absolute error (MAE) are regularly employed in model evaluation studies. Research from Willmott and

Matsuura (2005) have implied that RMSE is not a good indicator of average model performance and that it is a misleading indicator of average model performance. Moreover, they concluded that MAE would be a better metric to evaluate model performance. However, research by Chai & Draxler (2014) suggest that the MAE is not superior to the RMSE and that to measure a model's average performance at least a combination of the two should be used, preferably more.

In this paper we will use mean error (ME), mean percentage error (MPE) root mean squared errors (RMSE), mean absolute errors (MAE), and mean absolute percentage errors (MAPE) to measure the average statistical performance of our model. However, we will also test our model up against real values and a forecast with a simple linear regression to check the forecasted values against actual values. This to ensure that we are not misled by the statistical measurements.

4. Data

There is a substantial amount of internal data, therefore, we have decided to split the chapter into two main parts to better explain the data. The first part revolves around the internal data where we have 14 230 000 observations in our data set, where each observation has 37 variables. The second part revolves around the external data which consists of the rental market survey, the income and wealth statistics, recipients of disability benefit, and work clarification benefit (AAP) statistic.

4.1 Internal Data

Our internal data is supplied by the NSHB. Their data is collected from applications for- and recipients of housing allowance. The data set contains 14 230 000 observations. Each observation has 37 variables, including variables such as household ID, income, expenses, and age. There are 457 360 unique households, which means that many of the household have been in the system multiple times. The timespan of the data is from the 1.st of January 2010 to the 1.st of March 2020. In appendix b), a full list of the variables and descriptive statistics are available.

In our forecast and analysis of new recipients all data from 2010 are removed. The reason for removing 2010 data is that the formula used to calculate new recipients are biased for the first year. For instance, new recipients, every household are categorised as a new recipient in the first year as 2010 is the first year in our dataset. However, this is not completely true, due to the fact that the housing allowance practice started before 2010. This entails that a lot of people also received housing allowance prior to 2010. The reason that our data set does not date back longer than 2010 is because of a reform that was introduced in 2009. The previous reporting of these numbers had a completely different set up prior to 2010 and is therefore not comparable with the data collected after the reform. Furthermore, in the data used in the forecasting, we have decided to remove the data regarding households which has gotten their application declined. After the removal we are left with 12 930 000 observations and 356 530 unique households.

Moreover, in January 2017 the income basis in the calculation of housing allowance was changed to monthly tax income from “a-ordningen” from a yearly income basis. Prior to 2017 the housing allowance followed a clear pattern from month to month throughout the year. The number of recipients increased steadily throughout the year, followed by a significant reduction when the new tax assessment was made available in June. This means that the same household can receive housing benefit for months with low income, and not receive support in months with higher income. The latter also explains how the same household can be in the system multiple times. The transition to a monthly income base resulted in far greater variation from month to month. However, the change does not have a significant effect on new recipients. The effect was far greater to those who already received housing allowance. Thus, we chose to include data prior to 2017 in our forecasting of new recipients.

Lastly, for the purpose of cancelling out the effect of inflation this study looks at real rather than nominal values. It is worth noting that virtually all lease contracts written in Norway are adjusted annually for inflation via the consumer price index (CPI). This is important because 80% of the recipients live in rentals.

4.2 External Data

The external data used in this study is collected from Statistics Norway (SSB) and NAV. A brief description of the external data will follow in the following sections. For a thorough description, see appendix c) - h).

4.2.1 Rental Market Survey

The purpose of the survey is to measure rent levels in Norway grouped into different segments of the rental market. Although the survey dates back to 2005, it became an official statistic in 2006 and was further expanded and made more detailed in 2012 (SSB, 2019). We chose to include this data as the majority of recipients of Housing allowance are in the rental market. An increase/decrease in rental cost will directly affect their (approved) housing expenses. Thus, it will directly affect the size of the received allowance.

4.2.2 Income and wealth statistics

The income and wealth statistics by SSB provides figures for the level, composition, development and distribution of income and wealth. Furthermore, the statistics comprises all monetary income, both taxable and tax-exempt, as well as wealth and debt (SSB, 2019). Norway does not have an official definition of poverty in terms of income; therefore, the EU 60% definition is used. This entails that if a person does not earn at least 60% of the median, the person will be described as living in poverty (Skiphamn, 2020). We chose to include this statistic because it shows the level of people living in poverty and therefore it may help explain the gain or loss in the number of recipients of housing allowance. A high level of people/households with low income should result in more recipients of housing allowance and vice versa.

4.2.3 Recipients of disability benefit

The statistic shows both actual numbers and the share of the population who receives these benefits. It covers ages between 18-67. The numbers can be broken down to national, region and provinces (SSB, 2020). We have included this statistic for the purpose of explaining variations in the number of recipients, both future and historically.

4.2.4 Work clarification benefit (AAP) statistic

This statistic includes people who receives AAP. AAP is a benefit you receive from the government after you for various reasons are out of work. The statistic is updated monthly, new recipients and persons who no longer receive AAP is published quarterly (NAV, 2020). We chose to include this as it can help explain the development of housing allowance recipients. For instance, it is natural to assume that if more people are recipients of AAP, there would be more recipients of housing allowance.

4.2.5 Labour force survey

The labour force survey is the basis of our employment time series and unemployment time series. It is a monthly statistic that dates back to 2010, however we use figures from January 2011 – March 2020. The statistic is seasonally adjusted and are three-month moving averages. Changes are therefore calculated from figures published three months earlier (SSB, 2020). We chose to include these two time series in our forecast as employment and unemployment may help to explain the inflow of new recipients.

5. Methodology

As our research questions suggest this report is split between a description of who the recipients are and how the development have been in the period 2010-2020, and the prediction of inflow of new recipients. In this chapter we will expand on the methodology we employed to forecast the inflow of new recipients from 2020-2022 using a Vector Autoregression model on a multivariate time series. The model is estimated using the sample from January 2010 to March 2020, forecasting the 2 years out of-sample period until March 2022.

5.1 Choosing the forecasting technique

According to Chambers et. al. (1971), the selection of the method or technique depends on the context, the availability of historical data, the desired accuracy, timeframe for the forecast and the value of the forecast to the organisation. Given the data available in this case study, we will focus on time series analysis and forecasting. More specifically, multivariate time series analysis.

A multivariate time series has more than one time-dependent variable. Multivariate processes arise when several related time series processes are observed simultaneously. Meaning, that each variable not only depends on its past values, but also, on other variables. We use this dependency to forecast future values (Singh, 2020).

5.2 Vector Autoregression (VAR)

A VAR model is a systems regression model with more than one dependent variable. Hence, it lets us predict multiple time series variables using a single model. The VAR model extends the idea of univariate autoregression (Pexeiro, 2019). Meaning, the values of each of the g variables in the system depend on k lags of values of the other variables and error terms. Contrary to a linear regression model where all the variables are exogenous, all the variables are treated as endogenous.

The inflow of new recipients is forecasted based on four housing allowance specific variables and two macroeconomic variables. The first four variables are the number of new recipients (NEW), Average Housing Expenses (AHE), Age

(AGE), and Regulations (REG). The two macroeconomic variables are Employment (EMP) and Unemployment (UMP). Based on fundamental analysis principles we expect our time series to be non-stationary. Thus, we use an augmented Dickey Fuller (ADF) test to test for non-stationarity. In order to avoid an unpredictable forecast and spurious regression we perform first-differences on the series that contain unit roots.

Based on conversations with the NSHB and the descriptive analysis of data from January 2010 – March 2020, we do not expect the changes in our variables to have instant effect. Therefore, we use lags. The optimal number of lags to include are chosen through the LAGSELECT() function in R.

When performing diagnostics on our model several tests are used. It is tested for heteroskedasticity using ARCH test. Tests of normality will be performed by applying the Jarque-Brera (JB) test. Testing for structural breaks is done using CUSUM. To test for seasonality, we will first plot the distribution before we will deploy the WO-test, developed by Webel and Ollech. To test for serial correlation, we use an asymptotic Portmanteau test. Granger causality tests are applied to check for joint significance of all lags of the variables. The FEVD method in R is used to test the influence of each data series.

Finally, R is used to produce a forecast of the inflow of new recipients in the period 1st of April 2020 – 31st of December 2021. The forecast includes trend and seasonality.

5.3 Performance Evaluation

To determine the precision of the forecast we will use several measurements of performance. The measurements are, ME (mean error), RMSE (root mean squared errors), MPE (mean percentage error), MAE (mean average error), and MAPE (mean average percentage error). It is desirable that these scores are as close to zero as possible. Additionally, we have split the data into a training-and test set. Where we excluded the last twelve months in the training set. The reason for doing so is that we than can compare the predicted forecast with actual values. We have also made a basic linear regression, so we can see how well the VAR model performs compared to the basic linear regression model.

6. Analysis & Results

In this chapter we perform an empirical analysis of the recipients of housing allowance the past ten years and implement our VAR model. There are two main parts to this chapter, part one consists of chapter 6.1, 6.2, 6.3, and 6.4. This is the empirical analysis, here we will describe the development of housing allowance applicants and recipients. Moreover, we will use external data to compare this development up against the development of the Norwegian population in general. In part two, we will predict the inflow of new recipients in the period 1st of April 2020 – 31st of December 2021. Additionally, we will look at how extreme changes in both internal- and external factors affects the inflow of new recipients.

6.1 Grouping

We will be using the same groupings as the NSHB. There are three forms of grouping: geographical groups, user groups, and disposal form. The grouping applies for applicants and recipients.

6.1.1 Geography

To analyse the geographical differences between applicants and recipients we will use the same municipality groups that the NSHB divides their applicants and recipients in. Municipality group 4 consists of every municipality that is not included in group 1-3. Norway has 356 municipalities, which means that group 4 consists of 341 municipalities.

Group	Municipalities
1	Oslo
2	Bergen, Bærum, Stavanger, Trondheim, Tromsø
3	Frogn, Nesodden, Oppegård, Asker, Lørenskog, Skedsmo, Kristiansand, Sandes, Sola
4	341 municipalities

Table 1: Municipality groups

6.1.2 User group

The housing allowance recipients are divided into 5 separate categories. The determinant for which group they are placed in is source of income. In the table below, an overview of their main income source is included.

User Groups and their income		
#	Group	Income
1	Young Disabled	-Disability aid with status and rights as young diabled
2	Otherwise Disabled	-Disability aid
3	Elderly	-Retirement benefit (pension) -Financial assistance for elderly imigrants
4	Houshold with temporary benefits	- Work clarification benefit - Introduction benefit - Qualification Benefit - Transitional Benefit - Survivors pensions - Unemployment Benefit
5	Houshold without benefits	- Earned income - Other income

Table 2: User group sorted by their source of income

6.1.3 Disposal form

There are 9 categories of disposal form. The recipients and applicants are divided into groups based on their residential status.

Disposal form		
Group	Sub-category	Explenation
Public shared accomodation	Renting	Shared accomodation owned by the municipality
Public housing	Renting	Rented from the municipality
Private shared accomodation	Renting	Shared accomodation owned by a private landlord
Private	Renting	Rented from a private landlord
Freehold apartment	Owning	The household owns the apartment
Condominium	Owning	Building in which each unit is owned by an individual, but the grounds is owned jointly
Housing cooperative	Owning	The company is owned by those who live in the cooperative, the partholders.
Bond apartment	Owning	The apartment is the bond
Stock apartment	Owning	The apartment is a stock option

Table 3: Grouping based on disposal form

6.2 Housing allowance applications 2010 – 2020

We will start our analysis with a closer look at the total number of applications that the NSHB has received in the period January 2010 and until our last data point in March 2020. The highest number of received applications from unique households in a year was in 2011 with 181.248 applications. This means if a household has applied more than once, only one application is counted. The highest number of received application in a month was July 2011, with 134.685 applications from different households. Since July 2011, there has been a steady decline in applicants, cumulating in the lowest number of applicants in February 2020. A reduction of 24%. The general figure below shows the trend from 1st of January 2010 – 1st of March 2020.

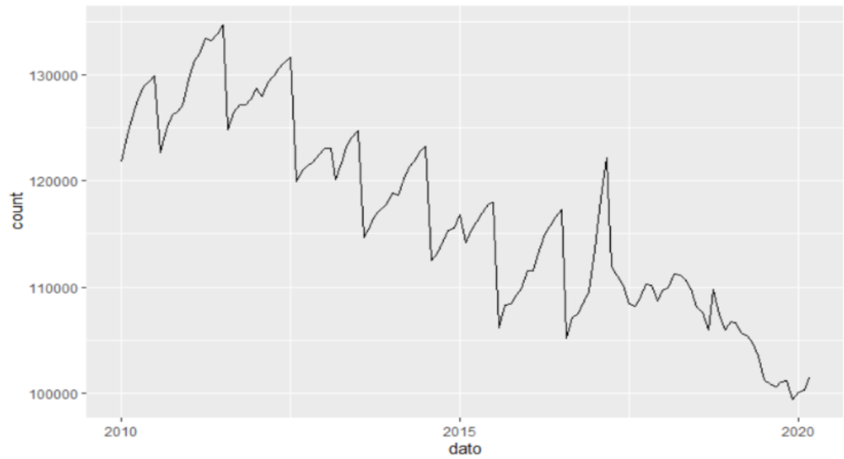


Figure 1: Total number of applications, development 2010-2020. The Y axis shows the count of unique applications. Furthermore, to show the development, the first value at the Y- is 900.000. On the X-axis one can see months.

When examining the development more closely, one can see that we are experiencing seasonality every July from 2010 – 2016. This is due to that every year prior to 2017 the income was calculated yearly, and in June there was a new income basis available. This led to that many applicants got rejected in June. Moreover, the trend shows a somehow linearly increase from August until the next July. In 2017, the income basis shifted from a yearly income basis to a monthly income basis. After this shift we experience a new trend – the yearly drops in July are now eliminated. In March 2017 there was an increase in applications before it dropped significantly. Except for the abnormality in March there has been a somewhat constant downward development.

From the general trend we have seen that there has been a negative trend in number of applications and therefore applicants. However, it does not provide us with any information about who the applicants are. To get a better understanding of the applicants we will look closer at their attributes and how these have developed over time. We will do this by looking at where they live, how they live, how old they are, and what income they have.

6.2.1 Municipality group

From the figure below we can see that most of the applicants live in Municipality Group 4 and that this has been the case throughout our time frame. In March 2020 Group 4 is 33% bigger than Group 1, which is the second largest group. As you saw in chapter 6.1, municipality group 4 is the biggest group in number of municipalities and inhabitants, so this is not surprising. However, whereas the 3 other groups seem to have a stable trend, Group 4 is experiencing a downward trend. From the top in 2011 until March 2020 there has been a decrease of 36% from 87.640 applicants to 56.157 applicants.

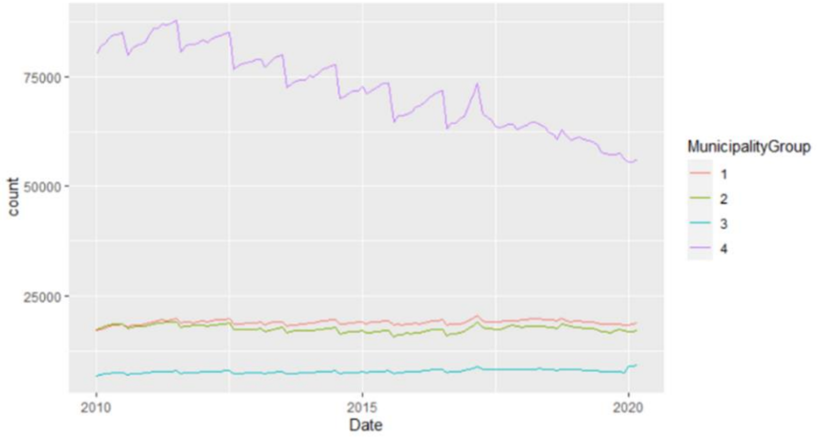


Figure 2: Development in total number of applications by municipality group 1-4, 2011-2020. The Y-axis contains the total number of unique applications per household per municipality group. The X-axis shows year. As can be seen from the figure, Group 4 is the largest group by application count.

Because group 4 is so dominating in terms of number of applicants it is hard to see the development of the other groups. In appendix A), a figure excluded group 4 can be viewed. When examining groups 1-3, one can see that group 2 was marginally larger than group 1 in 2010. However, from August 2010 group 1 has been bigger. Since this shift, group 1-3 have experienced a similar development up until 2020. In 2020 group 1 and 3 have increased 10% and 33% respectively, whereas group 2 have dropped 2%.

6.2.2 Disposal form

In this next part we will take a closer look into which disposal form the applicants live. From the figure below we see that the two dominant forms are privately rented housing and public housing. In 2010, most applicants lived in public housing whereas in 2020 privately rented housing was the most common. The two disposal forms have had a very different development, and in March 2020 the gap between applicants who lived in privately rented housing and public housing was 52%.

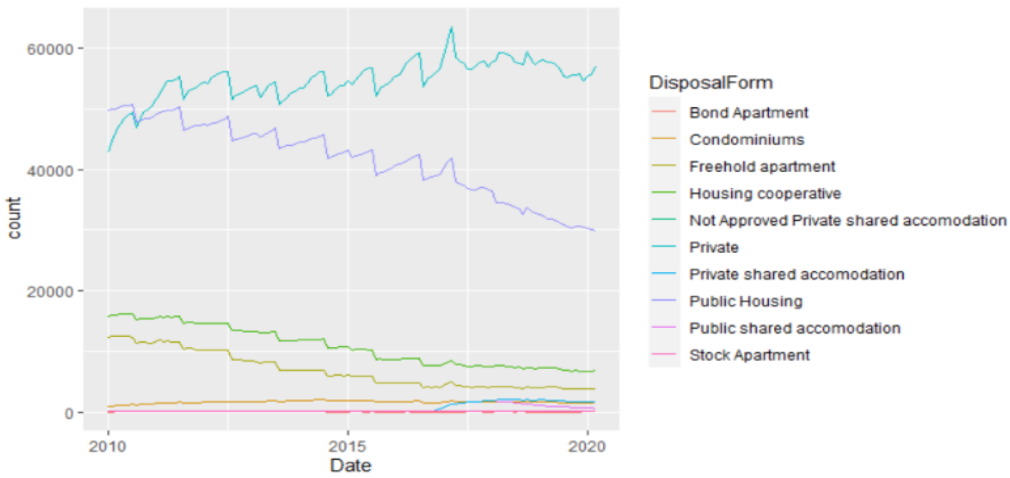


Figure 3: Development of applicants based on disposal form, 2010-2020. Recipients are on the Y-axis, and months are on the X-axis. The two dominant forms are privately rented housing and public housing.

With the figure above we face similar issues as we did regarding which municipality group applicants lived in. Privately rented housing and public housing are so dominant that it is hard to see anything else than a stable downward trend for the rest of the disposal forms. Therefore, in appendix A), a figure excluding the two disposal forms can be viewed. When examining the other disposal forms, we find that the two disposal forms which are decreasing are both disposal forms involving the applicant owning his or her apartment/house. More specifically, applicants living in a housing cooperative and applicants living in a freehold apartment. These two have decreased with 57% and 68 % January 2010 until March 2020. Another interesting detail which becomes clear by examining the data and can be seen in the aforementioned figure in the appendix is there was created a new disposal form in 2017, private shared accommodation. From its creation it increased from 246 applicants to 1600 applicants in 2018, before it stabilised. Lastly, applicants living in condominiums have increased with 70% from January 2010 until March 2020.

6.2.3 User groups

As described in chapter 6.1, applicants are categorised in 5 different groups based on their income. To obtain a better picture of who the applicants are and who the applicants have been in the period 2010-2020 we have plotted them in a time series. From the figure below it becomes clear the different groups have experienced a different development. Group 5, households without benefits, is the biggest group and has been throughout the period, except for a short time in the end of 2010 and right after the change in 2017. In March 2020 there were 39865 applicants from Group 5. Moreover, group 5 has had a relatively stable increase from 2011 until the sudden drop of 25% in 2017. The sudden drop of group 5 and the sudden increase in group 4, people with temporary benefits is caused by a change in how the NSHB categorise their applicants. The change in categorisation involved people that received work clarification beneficiaries was moved from Group 5 to group 4. This recategorization has contributed to Group 5 and Group 4 being significantly bigger than the other groups. Another interesting development is the development of group 3, elderly. In 2010 elderly people were the second biggest group of applicants with 30 000 applicants whereas in 2020 it was 12 811 applicants, a drop of 57%. Furthermore, we can see that group 1, young disabled, and group 2, otherwise disabled, both have decreased with 65% and 24% respectively.

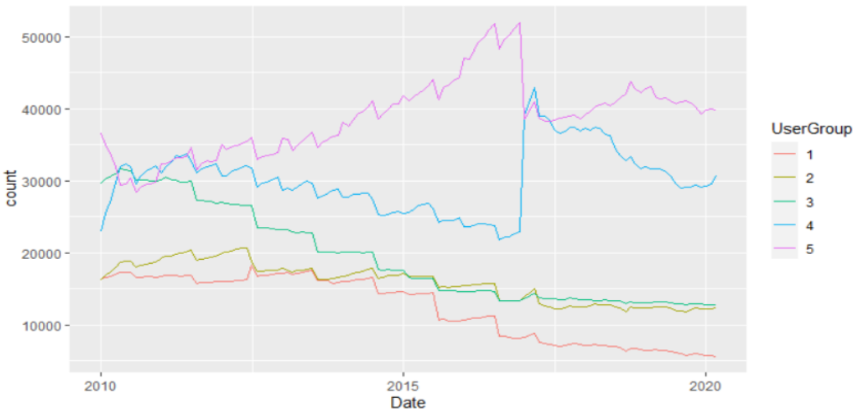


Figure 4: Development in applicants from different user groups, 2010-2020

From the section and figure above, we saw that applicants categorised as elderly had decreased with 57% This is evident when we look at the development of applicant age on a yearly basis. The average and median age for an applicant was 50,66 and 46 in 2010 whereas 43,92 and 40 in 2019. We chose to do the comparison with 2019, due to insufficient data for 2020, as we only have data from January-March. In appendix A) these data are visualised using a box-plot.

6.2.4 Income distribution

In this part we will take a closer look on the applicants' income. Even though this is not the recipients, we remember the hypothesis that future recipients would have an income that was close to nothing. The assumption from Fjelltoft & Ezat (2019) was based on the development of recipients from February 2010 – February 2019. If we apply this hypothesis to applicants to, that applicants who apply for housing allowance in 2020 will have an income which is lower than previously, it seems to have some traction. Although, the median income has increased, the distance between 1. Quartile and the 3. Quartile has increased significantly. If we compare the 1. Quartile in 2010 which was 121 494 to the 1. Quartile in 2019 which was 21 600, we see a decrease in income of 82% The reason we are not focusing more on the development of the mean and median income is that many of these applicants will not be granted housing allowance.

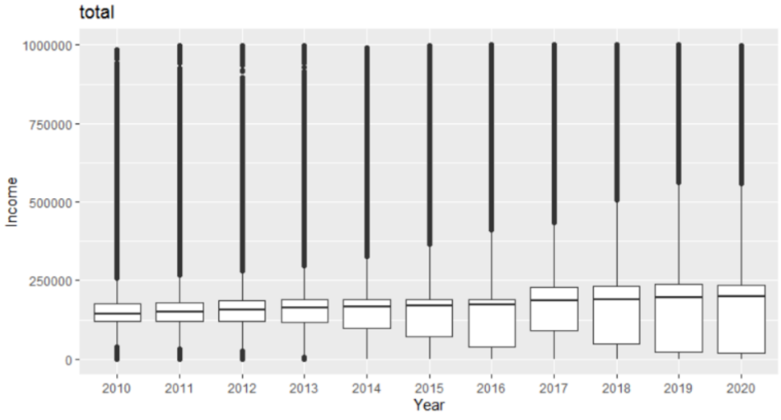


Figure 5: Income Distribution for total applications, yearly aggregated, 2010-2020

6.2.5 External data

In this last part of the analysis of total number of applications we will look at the external data to see if this could help explain the downward trend. Initially we found the downward trend somewhat surprising considering that our external data shows that rental expenses has increased by approximately 43% from 2012 – 2019, and this is the biggest expenditure for the average recipient of housing allowance. Furthermore, the number of people considered poor in Norway has increased from 2010 – 2017 (SSB, 2019). Moreover, the number of people on disability benefits has increased by 17% from 2011 - 2020. These factors would all suggests that the number of people that applied for housing allowance would be increasing. However, there are some findings that help explain the downward trend. Firstly, the number of people on Work clarification benefit (Arbeids Avklarings Penger(AAP)) has decreased by 30% from 2012 – 2020. And although the number of people on disability benefits has increased by 17% from 2011 – 2020, the number of people with disability benefits considered poor was just 12% (Amundsen, 2019). Lastly, the number of pensioners that was considered poor in 2008 was 17,5 %, whereas it was reduced to 9,3% in 2017 (Amundsen, 2019).

6.3 Declined housing allowance applications 2010 – 2020

The number of applications that has been declined has almost doubled since the beginning of 2010 to March 2020. If we look at the yearly aggregated data. We see that in 2010 there were 58 795 rejections whereas there were 99 402 rejections in 2019. The main reason for the significant increase in rejections is the previously mentioned change in income calculation in 2017. If we look at the figure below which shows the monthly development of rejections, we see that prior to the change we experience the same seasonality as we did in the total number of applications earlier in this chapter. It peaks in July and stabilises throughout the year. However, after the change we experience a much higher monthly variation and it changes much more from month to month.

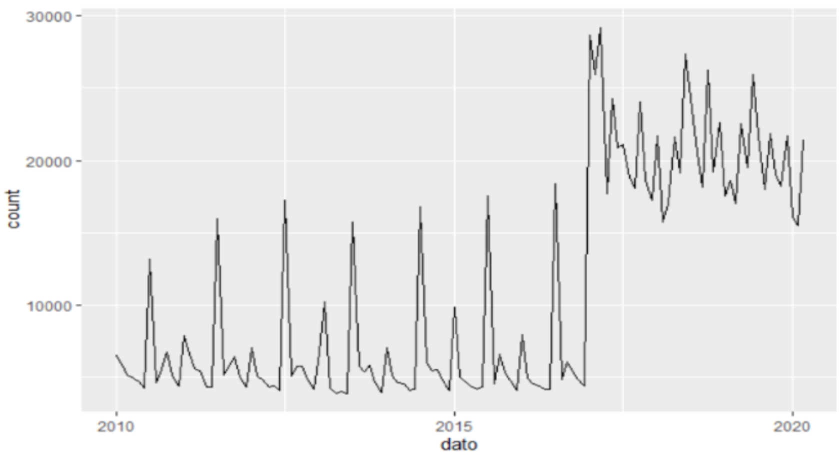


Figure 6: Total number of declined housing allowance recipients, 2010-2020. The effects of the change in income basis is instantaneous in January 2017.

6.3.1 Effects of the new income basis in 2017

When studying the data from 2017-2020, it becomes clear that the general trend applies for every municipality group. Hence, the change of 2017 affected the whole country equally. Moreover, it appears that the change does not affect one disposal form more than the others. Figures of visualising the effect of the change in income basis from 2017-2020 can be viewed in appendix A). Even though the change in 2017 affected the whole country equally geographically and in terms of disposal form, it affected some user groups more than others. These groups are user group 4 and 5. This becomes evident from the graph below. The reason for this is that households in group 4 and 5 is households with temporary benefits and without benefits. Their income changes from month to month, meaning that they will qualify for housing allowance some months whereas they will not qualify in others.

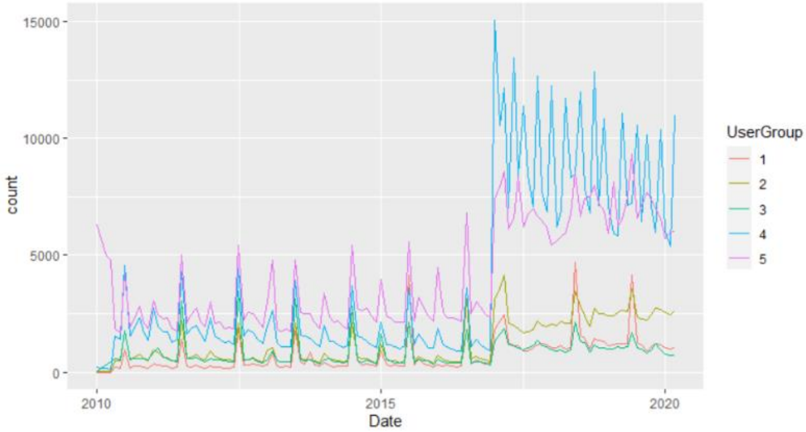


Figure 7: Rejection by user group, development 2010-2020. Recipients are on the Y-axis and months are on the X-axis.

6.4 Housing Allowance recipients 2010-2020

In this section of the analysis we will look closer into the recipients of housing allowance between January 2010 – March 2020. Overall, we see a negative trend. For instance, it was 115 196 households that received housing allowance in January 2010, whereas in January 2020 there was only 84 013 households that received housing allowance. Which is a significant lower number of households. The highest recorded number of recipients in our data set was June 2011 with 129 506 household received housing allowance. This means that since June 2011 there has been a 34% drop in recipients of housing allowance. The figure below shows the development from January 2010 – March 2020. When we study the development over time and compare it to the development of the total number of applications, we can detect a similar trend, which is reasonable since most of the applications gets approved. However, the yearly drop prior to the change in calculation of income in 2017 comes 1 month earlier for approved applications. Furthermore, after 2017 we can detect a new pattern with a higher monthly variation, but the overall variation is not as significant as before the change.

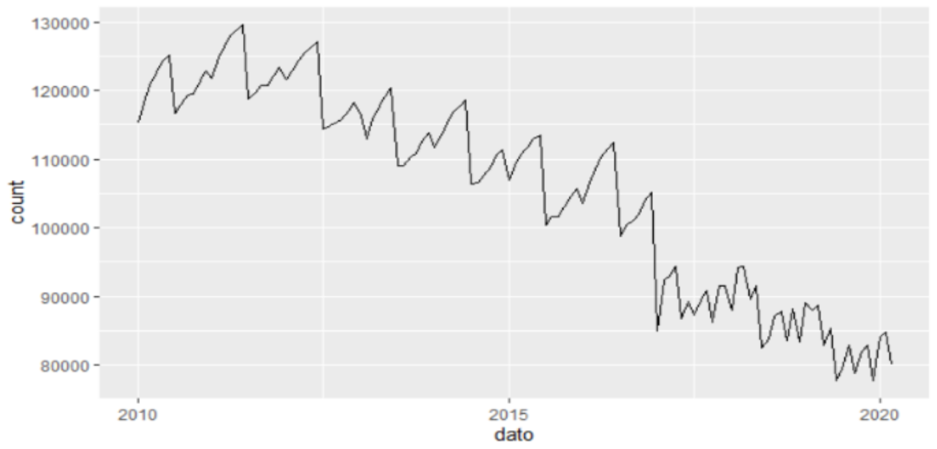


Figure 8: Housing allowance recipients, development 2010-2020. Recipients are on the Y-axis which starts at 75.000. Months are on the X-axis.

From the general trend we have seen that there has been a negative trend in number of recipients and that this trend is like the trend we saw regarding applicants. To learn more about the recipients we will analyse the recipients more closely. We do this by looking at their attributes and how these have developed over time. Specifically, by looking at where they live, how they live, how old they are, the household composition and what income, expenses and received amount they have.

6.4.1 Municipality Groups

We start by looking at their geographical location. Municipality group 4 is the biggest group. This is natural as they have the most inhabitants and the most applicants. From the figure below, we can see that group for has had a constant downward trend. In percentage, group 4 has decreased by more than 42%. Some of this drop can be explained by the change in calculation of the income. In 2017, group 4 experience a significant drop. In December 2016 the number of recipients were 105 207, whereas in February 2017 the number of recipients were 99402.This means that there was a 12% decrease in recipients in just 2 months.

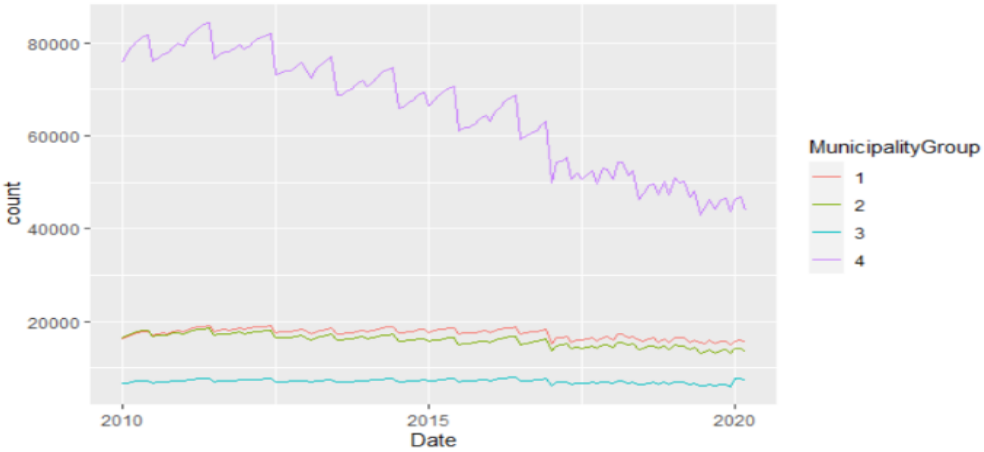


Figure 9: Monthly development of recipients by Municipality group 1-4, 2010-2020

Because group 4 is so dominant in the plot above it is hard to see how the development of the other groups have been. Therefore, in Appendix A), a figure showing the development of municipality group 1-3 is added. Groups 1-3 have experienced a more constant trend during the timespan than group 4. From 2010 - 2020 group 1 decreased 5%, group 2 decreased 18% whereas group 3 increased 11%. Although all the groups were affected by the change in 2017, groups 1-3 were able to “recover” quickly. And went up to the same level as before the change. Which means that the gap in recipients between municipality group 4 and 1-3 have decreased drastically in 2020 compared to the numbers from 2010.

6.4.2 User Groups

In this section we will look at the development of recipients in the different user groups. The first thing that becomes evident from the figure below is that group 4 and 5 is the two biggest groups. Group 5 has increased from January 2010 – March 2020 with 12% and is the only group which have not decreased in the period. Thus, it follows the same trend as we saw in total number of applicants which also experienced an upward trend. On the other hand, group 4, has 21% fewer recipients in 2020 than in 2010 despite an increase in the number of applicants. In the same way the change in categorization affected the number of applications for group 4 and 5 it affects the number of recipients in these two groups. Hence, it is the cause for the significant drop in user group 5 and the significant increase in group 4 in 2017. The decline of group 3 has been constant throughout the period and has declined 60% from January 2010 – March 2020. On the other hand, group 1 and 2 was stable until June 2012, group 2 even increased in this between January 2010 and June 2012, before both decreased from June 2012 - March 2020 with 71% and 51%, respectively.

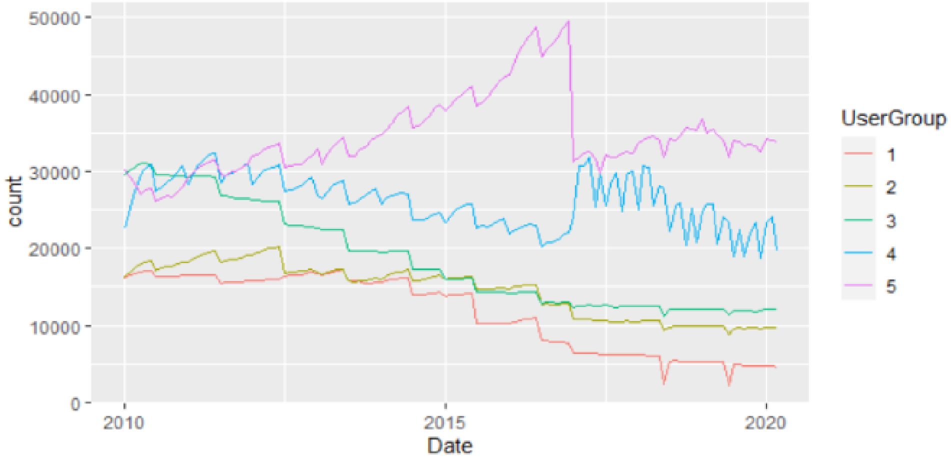


Figure 10: Housing allowance recipients on an aggregated level by group, 2010-2020

6.4.3 Disposal form

In this next part we will take a closer look into how the recipients live, meaning that we will first look at which type of housing form they live in. We will also look at what type of household they live, meaning how many people live in the household.

From the figure below we see that the two dominant forms are privately rented housing and public housing. The trend for these two categories is the same as the applicants' trend. In 2010, most of the recipients lived in public housing whereas in 2020 most of the recipients lives in privately rented housing. Privately rented housing is the only disposal form which has increased during our time-period. Following the trend of the applicants living form, the gap between who lived in privately rented housing and public housing was 40% which is slightly lower/higher than the applicants disposal form, were the gap was 52%.

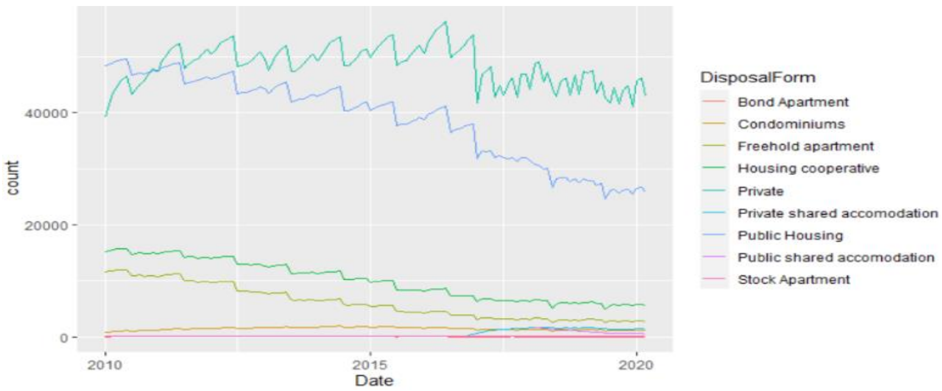


Figure 11:Monthly development of recipients living in different disposal forms, 2010-2020

Once again, we will remove the two biggest group to take a closer look at the other disposal forms. From the figure below, we see that all disposal forms follow the same trend as the applicants' disposal forms and that recipients who owns a freehold apartment and an apartment in a housing cooperative have decreased from 11 548 and 15 151 in January 2010 to 2 692 and 5 575 in March 2020. It must be mentioned that there are strict rules for wealth, this also applies for wealth in property, this makes it hard for recipients to get housing allowance if the tax value of the property is greater than the amount of mortgage they have left.

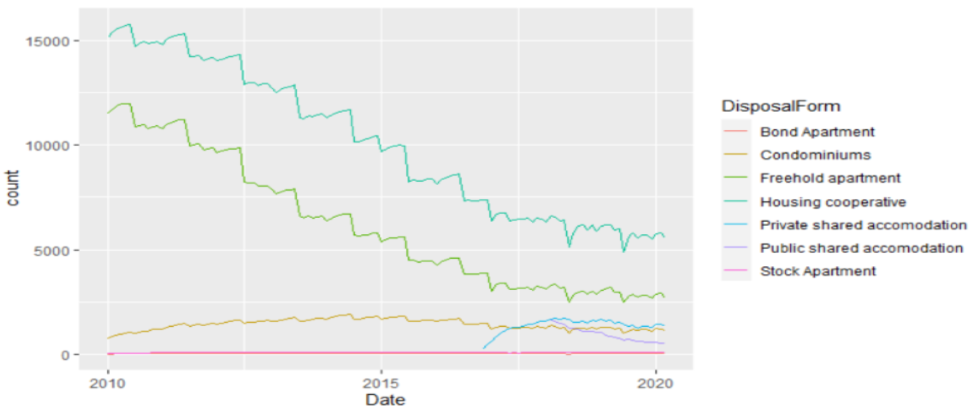


Figure 12: Monthly development of recipients living in different disposal forms, excluding privately rented housing and public housing, 2010-2020

6.4.4 Age distribution

As we know from the development in applicants and the development in user group the number of elderly recipients has decreased. The applicants average and median age were 43,92 and 40 in 2019. For recipients already in the system the average and median age is slightly higher at 45,20 and 41 in 2019. This is a decrease of 10% and 10,5% since 2010.

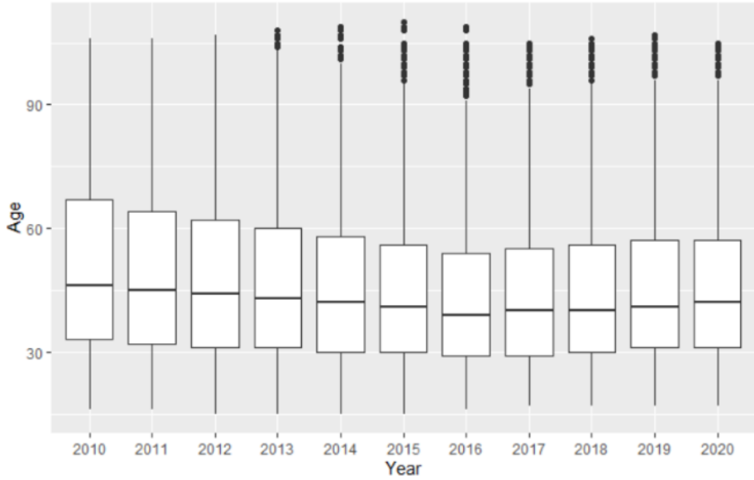


Figure 13: Development of Age for housing allowance recipients, 2010-2020

From the density plots below, we can see that the significant downward trend is mainly do the decrease of elderly recipients in Group 4. Since this group is the biggest the decrease would have been more significant than the others anyway, but as the plots show the age of recipients in the other three groups were more dense around a younger age than the recipients in group 4. The density plots also help explain why the average and median age among applicants are younger as it seems that many recipients that was in the system in 2010 are still in the system, thus making the population older.

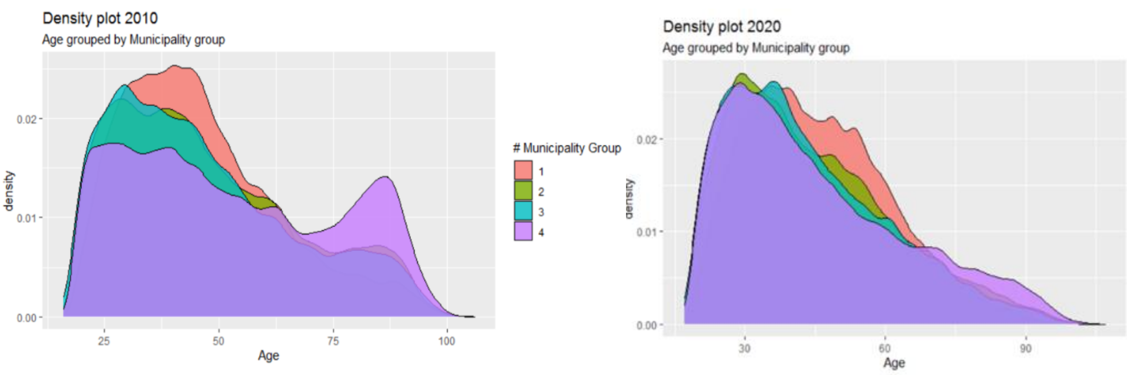
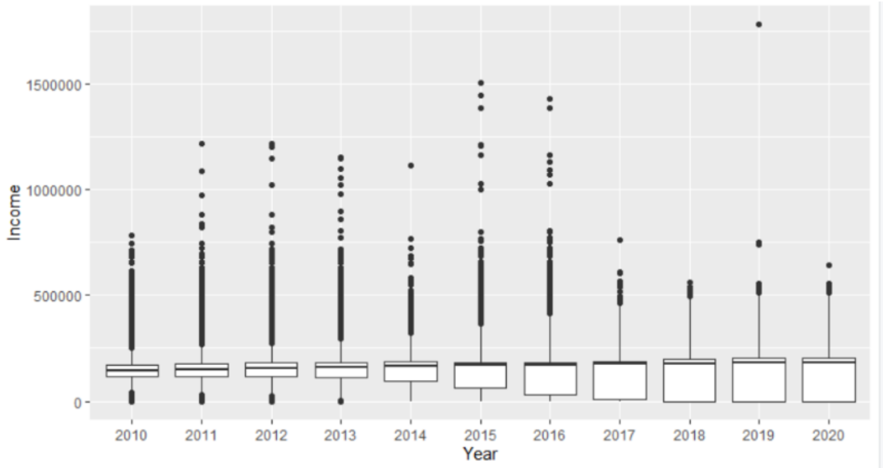


Figure 14:Age density by municipality groups. Left: Density plot 2010. Right: Density plot 2020

6.4.5 Income

When calculating how much a household will receive in housing allowance, income is an important factor. From the plot below, we see a clear trend that the first quartile is getting closer to zero. This is the same trend we saw regarding applicants and is what Fjelltoft & Ezat (2019) predicted in their report. The first quartile in 2010 was 119 201 whereas the first quartile was 0 in 2019. This confirms the assumptions made by Fjelltoft & Ezat (2019) regarding future development. If we look at the average income, this was the highest in 2013 at 140 352 and at its lowest in 2016 with 127 421. After this it is stable around 135 000. Even though the first quartile is now zero, the median has steadily increased throughout the period and was 184 320 in 2019, that is approximately 29% higher than it was in 2010.



When looking at the income trend on a monthly basis and by municipality group, we see that recipients living in Oslo (Group 1) has the lowest average income and that the gap has increased since the change in income basis in 2017. It is important to note that this is an average monthly trend showing the full year income.

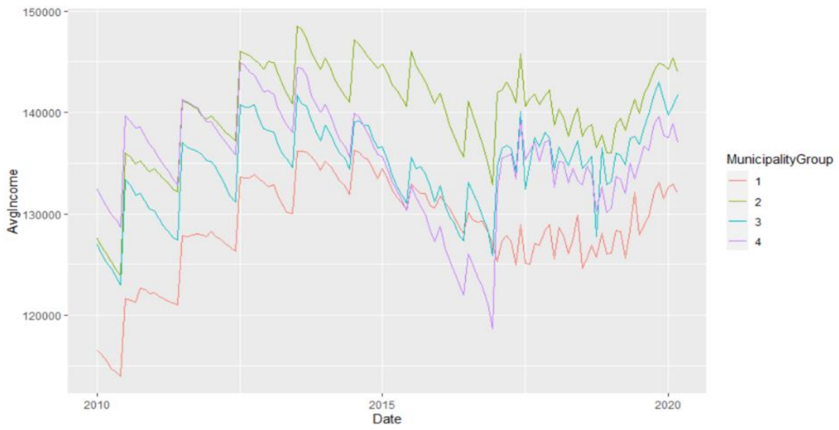


Figure 16: Average monthly income by municipality group on an aggregated level, 2010-2020

6.4.6 Housing Expenses

Like income, household expenses play a central role in the calculation of housing allowance each household receives. According to SSB rental prices in Norway has increased with 43% since 2010. Moreover, housing expenses is the biggest expenditure for any household in Norway (SSB, 2018). The increase in rental prices should indicate that housing expenses has increased similarly as approximately 80% of the recipients are in the rental market.

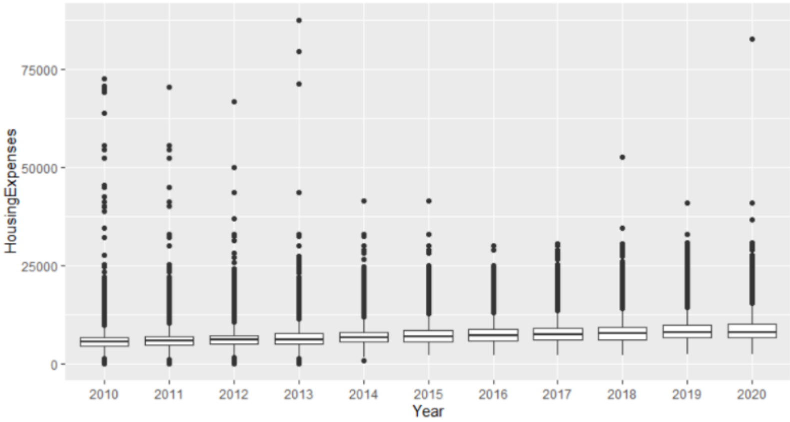


Figure 17: Development of yearly actual housing expenses for recipients of housing allowance, 2010-2020

From the figure above we can see that the lower quartile, the mean, and the upper quartile are all higher in 2019 compared to 2010. The lower quartile is 44% higher and is now 6 500, the upper quartile is 46% higher and is now 9 648, the median is 42% and is now 7 850. The average recipient's household expense has increased with 44% which is higher than the increase in rental rates alone.

NSHB operates with a maximum approved expenditure which is politically decided. If we compare the development of monthly approved housing expenses and the actual development of housing expenses in the period January 2010 – March 2020, we see that the approved housing expenses has increased with approximately 40% whereas the actual housing expenses has increased with 48%, Thus creating a gap of 8%. In the appendix figure A-figure 9 shows the development of approved housing expenses and actual housing expenses. The gap stems from that the NSHB only started to follow the rental market prices in 2017. Prior to 2017, the increase was politically decided in the national budget.

To further investigate the recipients expenses we split the expense data based on which municipality group the different recipients belongs to. As the municipality groups are sorted based on the price level of the different municipalities, it came as no surprise that Group 1 – Oslo – had the highest expenses as the rental prices in Oslo is higher compared to the rest of the country (NTB, 2019).

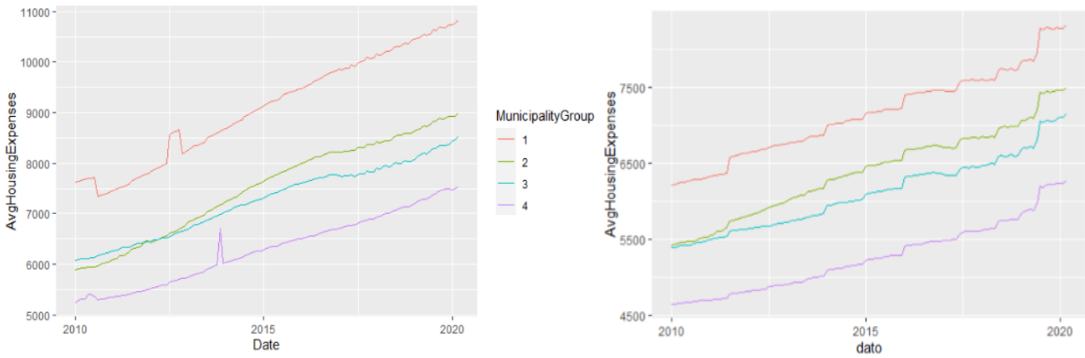


Figure 18: Development of actual average monthly housing expenses by municipality groups (left) vs development of approved average monthly housing expenses by municipality groups (right), 2010-2020

The housing expenses for people living in Group 1 is 10% higher than in Group 2, and 33% higher than the housing expenses for people in group 4. Another interesting find is the development of housing expenses in Group 2 and 3. In 2010, Group 3 had on average 3% higher actual housing expenses than group 2 whereas in 2020, group 2 had 5% higher actual housing expenses than group 3. However, if we take a closer look at the approved housing expenses for the same period it shows that group 2 had 0,7% higher approved expenses than group 3 in 2010 and 4,5% higher approved expenses in 2020. This discrepancy in the period from January 2010 to November 2011 between actual housing expenses and approved housing expenses is somewhat surprising given the fact that the municipalities were divided into groups based on housing expenses. This could indicate that the Ministry had a bias in their grouping of the municipalities assuming that municipalities with bigger cities automatically were more expensive than municipalities with smaller cities. The “bias” ended up being accurate as the actual expenses of group 2 eventually surpassed the actual expenses of group 3. This might imply that households in group 3 received insufficient housing allowance for almost two years. However, without further analysis it cannot be proved.

6.4.7 Received amount of housing allowance

Earlier in the study, we described the calculation behind the estimation of housing allowance. According to the calculation formula, Housing allowance = (approved housing expenses – deductible expenses) * Coverage percentage, the most influential parameter is housing expenses. There is an upper limit for housing expenses, this is based on the number of people in the household and in what municipality group you live in. Income is also an important factor when as will influence in what is characterised as deductible expenses. The higher income the higher deductibles. As more and more household as closing in on zero income, it becomes more and more redundant. From the box plot below we can see that there is an upward stable trend in the received amount of housing allowance. Since 2010 – 2019 the first quartile has increased with 3% from 1 435 to 1 482 the third quartile has increased with 36% from 2 772 to 3 772 and the median has increased with 27% from 2 134 to 2 725. This makes sense as the income has decreased and thus made the deductibles smaller and the housing expenses has increased.

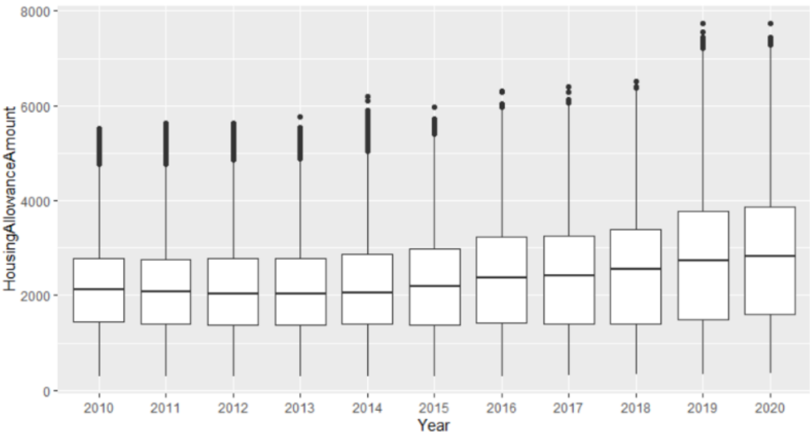


Figure 19: Yearly development in the received amount of housing allowance, 2010-2020

To further explain that housing expenses is a more dominant factor than income we have split the received amount of housing allowance into the municipality groups. The figure is monthly from January 2010 to March 2020 and shows average received amount of housing allowance. Group 1 has the highest amount of housing expenses and the lowest income and is the group that receives the most housing allowance. Which is natural. If we look at Group 2 and 3, we see the effect of housing expenses. In March 2020, Group 2 has 4,5% higher housing expenses than group 3 and 1,6 % higher income. However, they have 4,9% more in housing allowance than group 3.

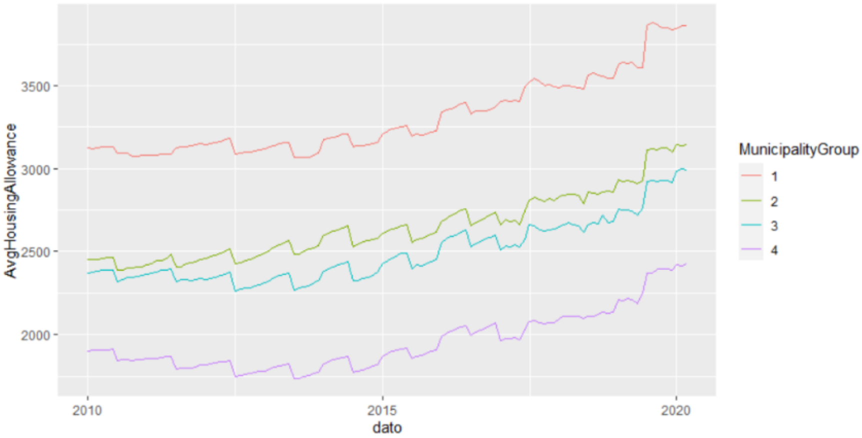


Figure 20: Development of average monthly amount of received housing allowance by municipality group, 2010-2020

6.4.8 Frequency of received housing allowance

In this part of the analysis, we take a closer look into how often the recipients between January 2010 and March 2020 have received housing allowance. The distribution is clearly right skewed which means that most households has received housing allowance in less than 40 months. The peaks in the distribution below is caused by the previous yearly income basis. There are 6 293 unique households that have received housing allowance each month throughout the period from January 2010 – March 2020.

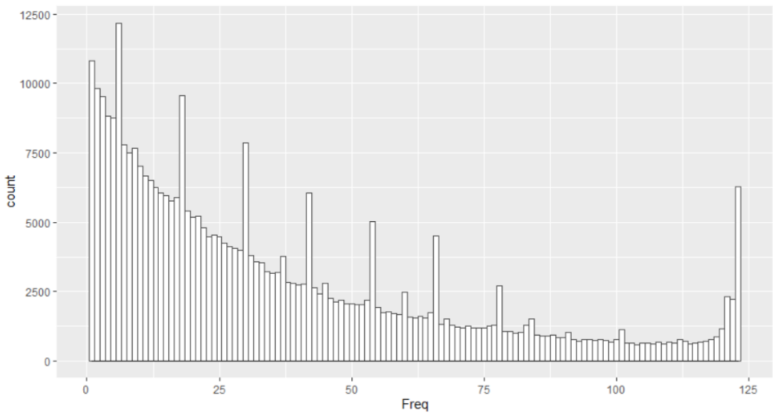


Figure 21: Distribution of number of months unique households has received housing allowance. Number of recipients are on the Y-axis and frequency are on the X-axis. Approximately 6000 recipients have received housing allowance every month from January 2010 – March 2020

To get a better understanding of the characteristic of households that have received housing allowance every month the past 10 year. We isolated them as a group. If we split the recipients based on their municipality group, we see that 50% of the recipients live in Group 4, 20 % lives in group 1, 20% lives in group 2, and 10% lives in group 3. This is quite similar to the rest of the recipients, which means that where the recipients live will not provide information about how long they will stay in the system.

On the other hand, the distribution of user group looks quite different. Households without and with temporary benefits are the two user groups with the lowest population, which is the opposite of what we could see for the total approved recipients. This is natural as their income will be unstable. However, as you can see from the figure A-figure 11 in the appendix, the biggest group that has received housing allowance every month throughout the period is group 3, elderly. Even though there are fewer applications from group 3, and fewer recipients of housing allowance from group 3. The density plot shows that recipients already in the system has gotten older, and therefore may have changed group as the years have passed. The user group otherwise disabled are the second most populated user group, which counts for approximately 20% of the total number.

Throughout this analysis we have seen that the change from yearly calculation to monthly calculation of the income basis has caused more variation in the number of recipients. Therefore, we wanted to see how many households that received housing allowance every month from January 2010 – December 2016. As you can see from the distribution below, this too is right skewed and has the same peaks as the distribution for the entire period. And more interestingly there are approximately 16.000 unique households which recieves housing allowance every month. Which means that after the change to monthly income calculation, only 6000 of these continued to receive housing allowance every month.

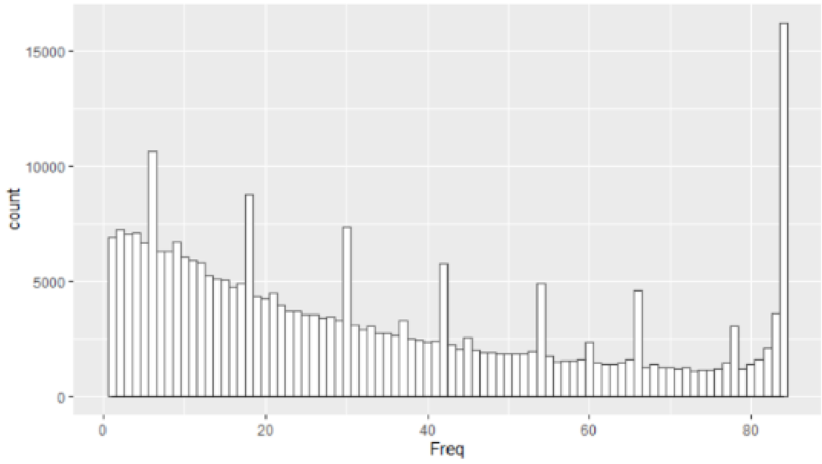


Figure 22: Recipients of Housing allowance distribution 2010-2016. Number of recipients are on the Y-axis and frequency are on the X-axis. There are approximately 16.000 unique households who have received housing allowance every month from January 2010 – December 2016.

6.5 New Recipients of housing allowance 2011 – 2020

Estimating new recipients is something that the NSHB has struggled to do accurately. The forecasting model they currently use do not give them an estimate of the inflow of new recipients, just a total number for a year. Before we embark on a forecasting model, we will analyse the distribution of new recipients from 2011-2020. By analysing the past inflow of new recipients, we hope to gain knowledge of who the new recipients are, witch characteristics they have. Moreover, we would like to find out which macro variables affect the inflow, and how the continuous changes in regulations affect them.

In the period from January 2011 to March 2020, there are in total 208 700 recipients who never have received housing allowance prior to 2010. It must be mentioned that there could be some bias in the year 2011. This is caused by that our data starts in 2010, meaning that people who did not receive housing allowance in 2010 but who may have received it prior to 2010 are categorized as new recipients. From the table below you can see an overview of the total population, percentage of the total population and the growth between 2011-2020.

Total population	2011	2020	% of the total population 2020	Growth from 2011-2020
Municipality group 1	599 230	693 494	12,9 %	15,7 %
Municipality group 2	740 927	837 371	15,6 %	13,0 %
Municipality group 3	366 826	472 763	8,8 %	28,9 %
Municipality group 4	3 213 322	3 363 952	62,7 %	4,7 %
Total	4 920 305	5 367 580		9,1 %

Table 4: Development in the general Norwegian population divided into the municipality groups used in this study, from 2011-2020

When analysing the general distribution of new recipients for the past 9 years, we can see a stable trend for the past 8 years. This becomes even more clear when studying the y-axis. As disclosed in the data chapter of this study, we have removed 2010 (the first year in our data set). Moreover, the recipients are considered as brand new, this means that they have never received housing allowance before, in our data set.

The figure shows that there is some variation over time and that there is a downward trend. The change in 2017, resulted in a spike in new recipients. However, the spike was short-lived, and the level quickly dropped back to the previously recorded level before it decreased further.

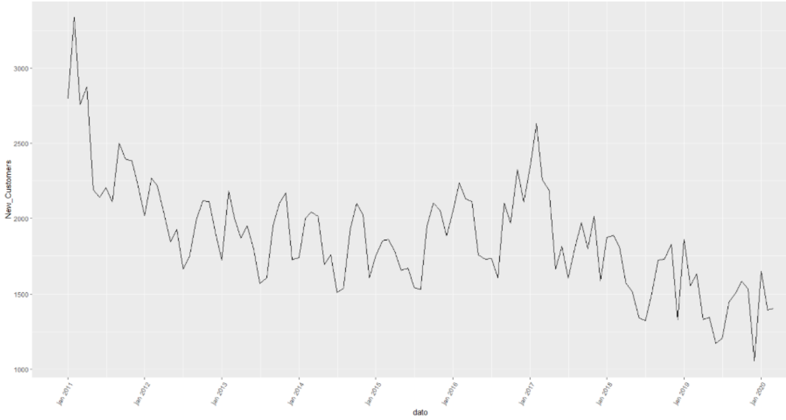


Figure 23: New recipients on an aggregated level, 2011-2020

6.5.1 Municipality group

When we take a closer look at the new recipients, we can see that approximately 50% of the new recipient lives in municipality group 4. This was also the case when we looked at the total number of housing allowance recipients' approvals. Municipality group 4 is also the group with the highest total population in Norway. Over 60% of the total population lives in municipality group 4, which means that they are underrepresented by 10% if we compare it with the total population in Norway in 2020. However, there have been a steady decline of new recipients in municipality group 4 since 2011. Despite the decrease in new recipients for group 4, the decrease is not that significant compared to decrease in the total number of approved recipients. After the change in 2017, it looks like the number of new recipients have "stabilized" at a lower level compared to prior the change.

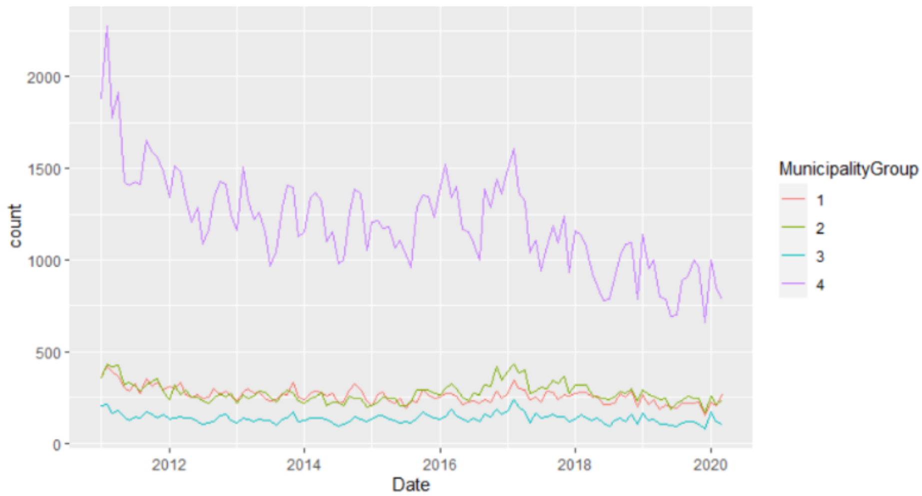


Figure 24: Development in new recipients by municipality group 1-4 from 2011-2020

To get a better look at the remaining municipality groups we will exclude group 4 in the next plot. Looking into municipality group 1-3, we can see an almost identical pattern as we could see for municipality group 4, just at a lower level. The ups and downs appear at the same time, which may indicate that there could be an overall trend these months. Municipality Group 3 had the lowest number of new recipients, group 3 is also the group with the lowest population in total. Prior to the change in 2017 there was a clearly upward trend of new recipients, especially for municipality group 2. After the change there was significant drop for a few months before it started stabilizing. Besides of the year 2011 and the months around the change, the number of new recipients for these municipality groups have been at the approximately same level, with some seasonal variation through time. Which may indicate that the decrease in total number of new recipients in mainly caused by municipality group 4.

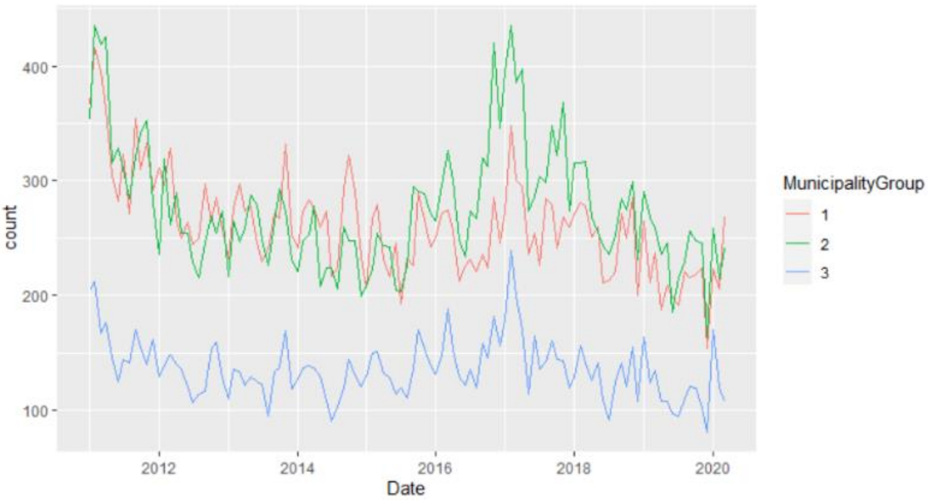


Figure 25: Development in new recipients by municipality group 1-3 from 2011-2020

6.5.2 Age

The average age for a new recipient was 34 years, and the median was 31. This is lower compared to the age of total recipients which had an average age of 45 and median age of 42. Meaning that the new recipients are more than 10 years younger, on average, than recipients already in the system. The trend for both new- and total recipients is that there are younger people who receive housing allowance, compared to 2011. This makes sense due to the inflow of younger recipients combined with the decreasing number of elderly recipients.

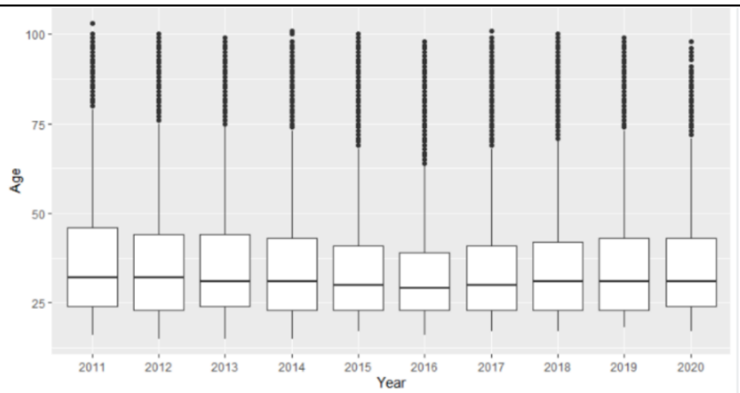


Figure 26: Age distribution for new recipients, yearly from 2011-2020

6.5.3 User groups

The new recipients mainly consist of households with temporary benefits and households without benefits, these two user groups were also the groups with the highest population for approved recipients. The distribution is very similar to the distribution of approved recipients. The only thing that we can detect immediately is that the user group, elderly, represent a smaller amount of the total number of new recipients. This may indicate that most of the elderly people have received housing allowance before, and there are not that many who starts receiving when they are categorized as elderly.

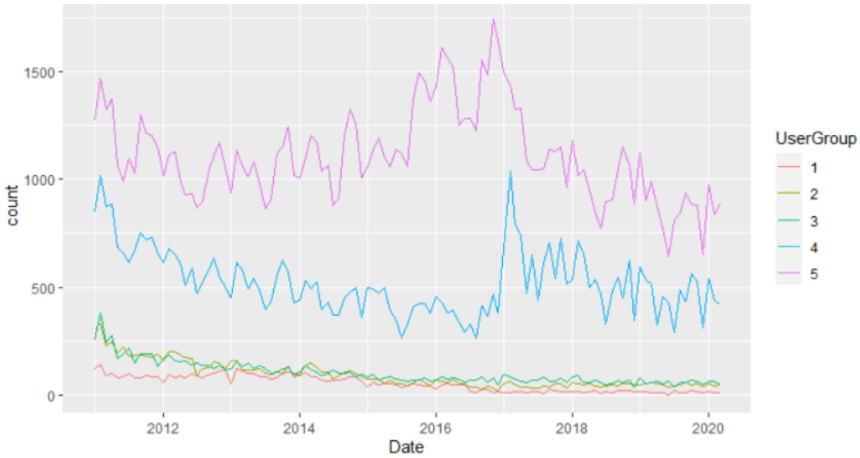


Figure 27: Development in new recipients by user groups (1-5) from 2011-2020

The report from SSB (2019) found that elderly have more money now than previously, which makes it reasonable that this user group have decreased from 2011-2020. For the disabled, both young and otherwise there have been a decrease in the number of new recipients - this is a really interesting finding, according to NAV(2020) there are more people in the Norwegian population that are disabled, both young and otherwise, compared to 2011. This should in practise mean that new recipients from these two groups should have increased since 2011, but from our distribution, this is clearly not the case. According to Fjelltoft & Ezat (2019), the main reason for this is the under-regulation of the rates used in the housing allowance calculation. Over time, the under-regulation will cause households who receives disability benefit to eventually exceed the maximum income limit and thus no longer qualify for housing allowance.

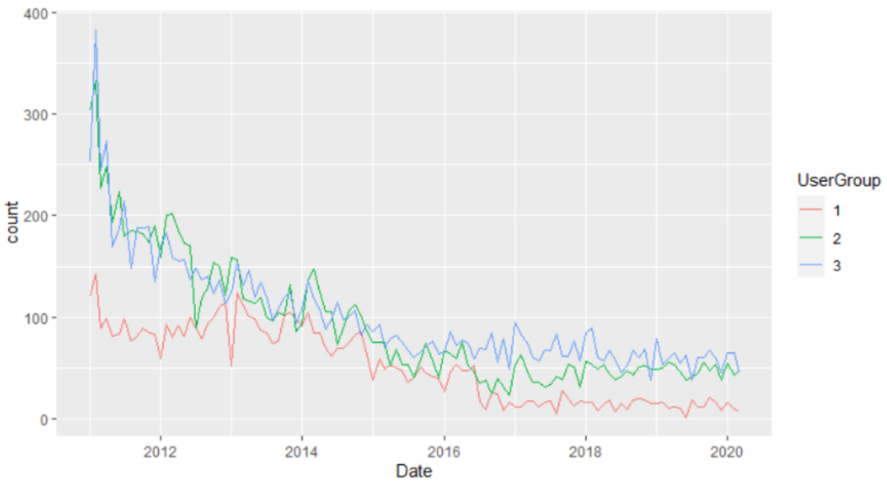


Figure 28: Development in new recipients in user group 1-3 from 2011-2020

6.5.4 Disposal form

Most of the new recipients lives in a privately rented housing, this was also was the case for the total number of recipients. The second largest group of new recipients lives in public housing. The biggest difference for new recipients compared to the total number of approved recipients is that recipients living in privately rented housing has been the biggest group during the entire time period. This could indicate that the trend we detected in the total number of approved recipients, may have been true also before 2010. However, both, privately rented housing and public housing has experienced a negative trend the past 9 years. Meaning that fewer new recipients live in either of the two disposal forms.

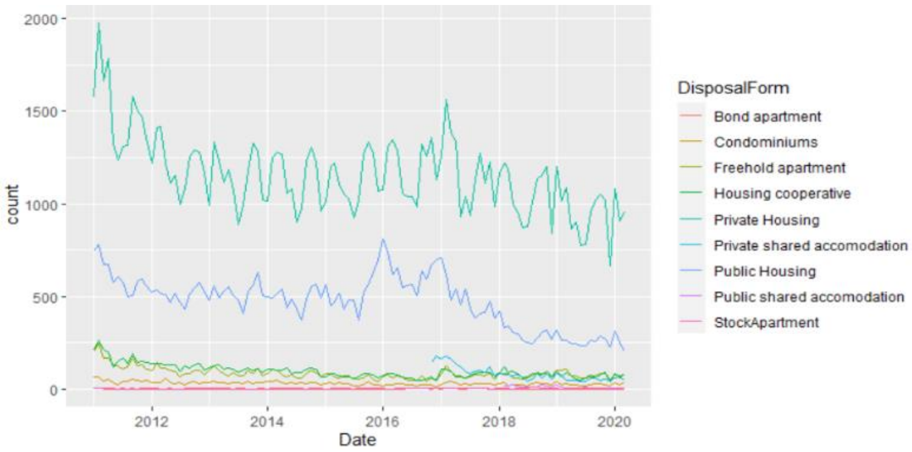


Figure 29: Development of new recipients by disposal form, 2011-2020

As you can see from the trend displayed above, privately rented housing and public housing are the two dominating groups of living accommodation. From the figure below, we can see that the rest of the disposals forms have had a steady devleopment since 2012, whith some seasonal differenceses. Housing coopereative and freehold apartment are the highest popullated forms, and they follow a very similar trend. We can aslo see that private shared accomodation was introduced as a category in 2017.

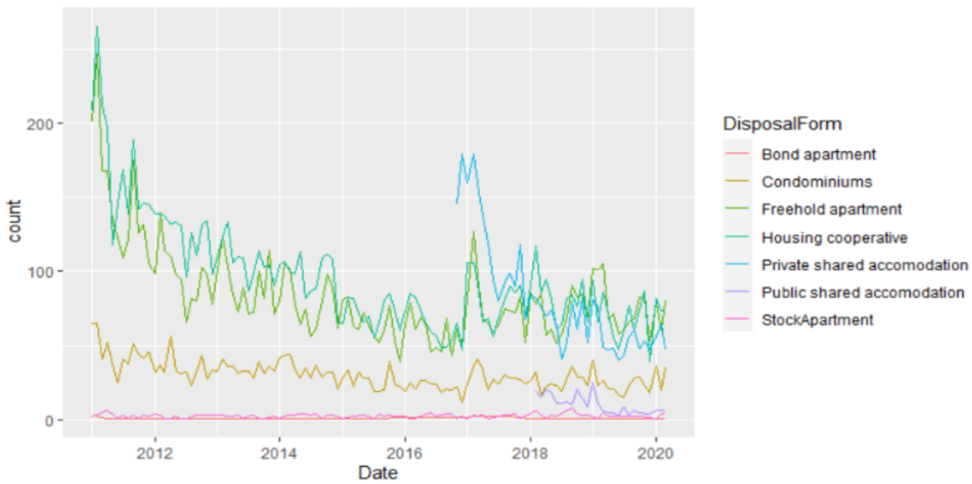


Figure 30: Development of new recipients by disposal form, excluding public housing and privately rented housing, 2011-2020

6.5.5 Income

The distribution of income from new recipients is significantly lower than what is the case for the total number of approved recipients. This indicates that most of the new recipients have a very low income, 1. Quartile value is 0 for every year since 2011. We can also see that the median value was lower in 2014-2016, whereas the median was as low as 464 NOK in 2016. This indicate that most of the values was close to zero that year. After 2016 the mean and the median of income have had an upward linearly trend, whereas the quartile is approximately at the same level. Compared to the total number of recipients, it can look as if the new recipients have less money available than the rest of the population. Compared to the median and average income in Norway, the income values for new recipients are extremely low. The income in only 15% of what the average earns in Norway. It has to be taken into consideration that housing allowance is calculated based on taxable income. Therefore, benefits such as child support, social assistance, basic benefit, and other similar benefits are not included. Hence, most households who receives housing allowance got more money than the income statement expresses.

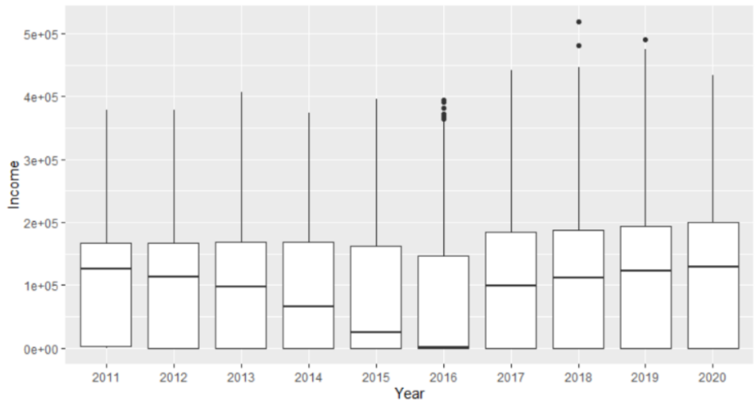


Figure 31: Average income distribution by year for new recipients, 2011-2020

6.5.6 Housing expenses

The average housing expenses is higher for new recipients compared to average expenses for total approved recipients. Furthermore, if we look at this in relation to income, which is lower for new recipients than for total approved recipients, it seems that new recipients are struggling more than existing recipients. One of the reasons that housing expenses is higher for new recipients than total number of recipients is the geographical differences in housing expenses. It is more expensive in group 1, Oslo, than in group 4 which consists of many small municipalities. Municipality group 4 is the biggest group in terms of total recipients and new recipients, however, as mentioned earlier they are underrepresented in terms of new recipients. This means that the effect of having the lowest housing expenses is smaller when looking at new recipients compared to what the effect is when looking at the total number of recipients. Overall, there have been an upward trend in housing expenses, which support the findings from SSB, that renting prices has increased with over 40% since 2011. A box plot of average housing expenses per year can be viewed in appendix A.d).

Another interesting finding is that average housing expenses for new recipients is only 5% lower than the average housing expenses in Norway. To put this into perspective the average income of new recipients is 80% lower than the average income.

6.5.7 Received amount of housing allowance

The income and total expenses for the new recipients was somewhat different from total recipients. However, the amount of housing allowance received are stunningly similar, the quartiles, median and mean are almost identical. Which may indicate that small differences in income and total expenses do not play as important role as we may think. One of the reasons for this is that a large share of the recipients has housing expenses that exceeds the approved housing expenses. Hence, fluctuations in housing expenses does not lead to increased/decreased amount of housing allowance.

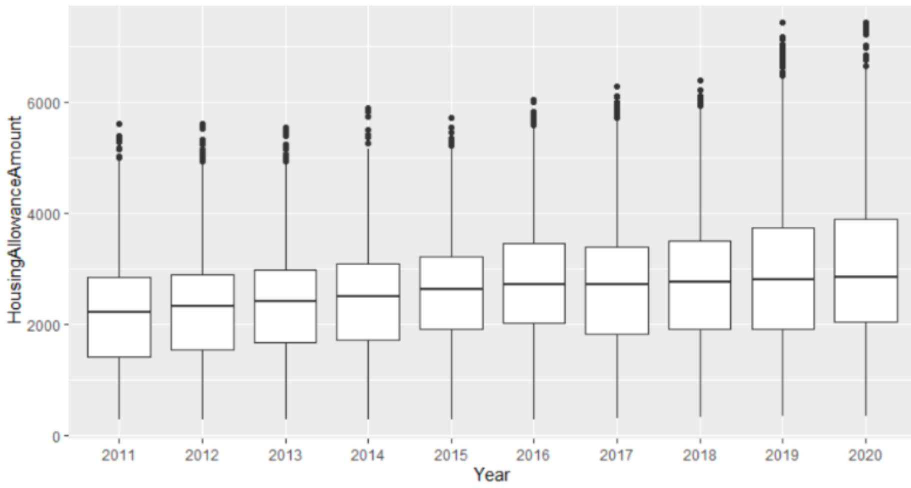


Figure 32: Average received housing allowance for new recipients by year, 2011-2020

6.6 Forecasting of new recipients of housing allowance

6.6.1 Model

We apply Vector Autoregression to investigate how the variables are affected by each other, when more than one variable changes over time. In addition, it is used to predict the monthly inflow of new recipients in the housing allowance scheme. The model explains the development of the endogenous variables as a function of the lagged values of them self and the other endogenous variables. We have decided to run the VAR model with 6 endogenous variables. The first one being the number of new recipients(NEW), the second is the average age of the new recipients, the third is the average housing expenses for the new recipients, the forth is the number of regulation, fifth is the number of employment(in 1000) and lastly we have unemployment. We are fully aware that there could be other macro variables that could be used. However, this was not possible as most of the macro variables are given on either a yearly or quarterly basis. Moreover, the data was constructed in such a way, that it was not possible for us to divide the data into a monthly statistic. Concerning the internal data included in the model, these were selected based on the analysis in the previous chapter.

To determine the number of lags to be included in the model we used the LAGSELECTION() command in R. The command will automatically calculate the preferred lag order based on the multivariate iterations of AIC, HWIC, SBIC and FPE. In addition, it is recommended to test various length of the lag, so we are certain that we have chosen the lag that fit our data best. Based on the automated calculation, Aikake’s info criterion suggested that the optimal number of lags were 10. However, when we ran the VAR model with 10 lags, we could detect serial correlation in the residuals and the forecast accuracy was low. Hence, we tried different number of lags and ended up with a lag of 3, which was chosen based on the FPE criterion.

6.6.2 Diagnostics

For the VAR model to be valid, the time series used must be stationary. In general, this means that the statistical properties are constant over time. Non-stationarity can lead to spurious regression. The variables were checked for unit roots using the Augmented Dickey Fuller (ADF) test. If the time series has a unit root, it is said to be non-stationary. For the ADF-test we have the following hypothesises, Ho: Serie contain unit root. H1: Series are stationary. The results from the test is showed in the table below.

	Original data	**First differenced**
Augmmentet Dickey-Fuller Test	**P-Value**	**P-Value**
NEW	0,07786	0,01
AHE	0,4765	0,01
AGE	0,9089	0,01
REG	0,01	0,01
UMP	0,9086	0,01
EMP	0,5992	0,01

Table 5: Results of the ADF-test on the original data and on the data after first differencing, In the original data only Regulations (REG) is stationary. After first differencing, every time series is stationary.

The results of the test indicate that all variables except Regulations contain unit roots, making them non-stationary. To make the non-stationary variables stationary, we transform the variables by first-differencing them. This creates monthly changes for each variable. Moreover, we first-differenced the regulations, even though the ADF-test showed that they were stationary to begin with. This was done to get a more consistent output. The output of the ADF-test after differencing. After the transformation all the variables are stationary. Thus, we can use VAR for our time series forecasting.

Test	**Degrees of freedom**	**Chi-squared**	**P-Value**
ARCH (multivariate)	5292	1743	1
Portmanteau (asymptotic	468	488,81	0,2445
Jarque-Bera (Normality)	12	141,46	< .2.2e-16
Jarque-Bera (Skewness)	6	27,659	0,0001089
Jarque-Bera (Kurtosis)	6	113,8	< -.2.2e-16

Table 6: Results of: ARCH (multivariate) test for heteroscedasticity, Asymptotic Portmanteau test for serial correlation, and Jarque-Bera (JB) (multivariate) test for normality, skewness, and kurtosis

To test for heteroscedasticity an Arch test is used. P-values close to or below the 5% threshold indicate heteroscedasticity. Our test result indicate that our model is homoscedastic. The result of the test can be viewed in the table 6. Moreover, to test for serial correlation, we use an asymptotic Portmanteau test. P-value of 0,2445 indicates that there is no serial correlation in in our residuals. Furthermore, it is also desirable that we have normality in the distribution of the residuals. To check for normality, we use a multivariate Jarque- Bera test. The test result indicate that our residual distribution is not normal distributed.

We also want to test the model for the presence of structural breaks. A structural break is an unexpected change over time in the parameters. This can lead to a drastic forecasting error, which again will lead to an unreliable forecast. To check for the presence of structural breaks we use a plot of the sum of recursive residuals. If the black graph, goes out of the red line, there are structural break at that point. As we can see from the plot, we do not experience any structural breaks.

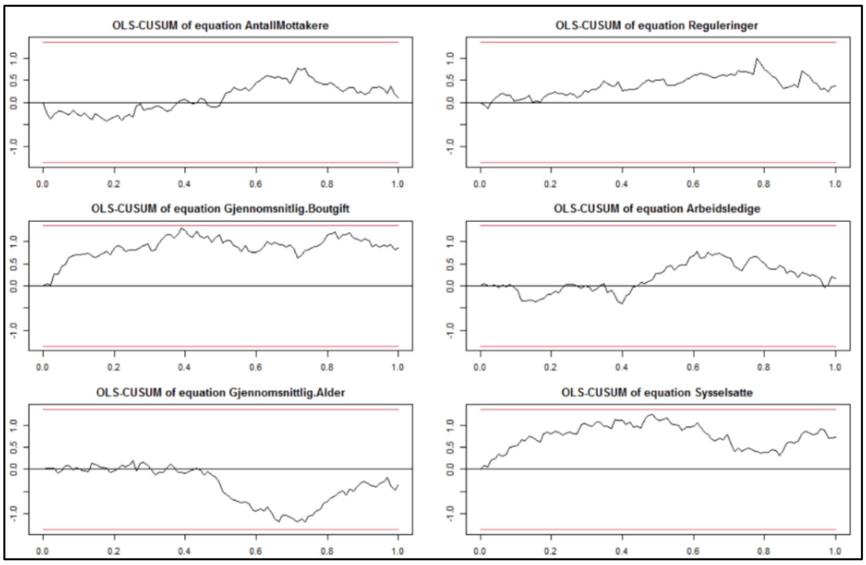


Table 7: Results of the sum of recursive residuals, test for structural breaks

6.6.3 Granger Causality

To check for causality, we employ a Granger causality test. Note that causality in this context does not mean that one variable directly causes movement in another, it simply suggests a chronological order of movements in the system. Some causality is found in the system for the estimation. The results can be viewed in the figures below. H0 is that the dependent variables do not Granger-cause the other variables, and do not have an instantaneous effect.

Dependent variable NEW

$Granger	F-test	df1	df2	P-value
	0.82178	15	390	0.6534

$Instant	Chi-squared	df	P-value
	20.934	5	0.0008335

Table 8: Granger test with New recipients of housing allowance (NEW) as dependent variable, the test shows that NEW does not Granger cause the other variables, but that there is an instantaneous effect

As you can see from the figure above, New recipients (NEW) does not Granger cause the other variables. This makes sense as new recipients will not have an effect on the external factors such as employment (EMP) and unemployment (UMP). However, we see that there is an instantaneous effect when we shorten our timespan. This is because the variables Age (AGE) and Average housing expenses (AHE) is calculated based on recipients, thus, new recipients will affect these values instantaneously.

Dependent variable AHE

$Granger	F-test	df1	df2	P-value
	3.6782	15	390	4.757e-06

$Instant	Chi-squared	df	P-value
	9.9582	5	0.07643

Table 9: Granger test with Average Housing Expenses (AHE) as dependent variable, the test shows that AHE does Granger cause the other variables, but that there is not an instantaneous effect

From the figure above, we can see that average housing expenses (AHE) does Granger cause the other variables. This may come as a surprise, since AHE is a product of the recipients. However, it is important to be aware that this is Granger causality, so even if x_1 does not cause x_2 it may still help pedict x_2 and thus Granger causes x_2. We do not detect instantaneous causality.

Dependent variable REG

$Granger	F-test	df1	df2	P-value
	2.1219	15	390	0.008517

$Instant	Chi-squared	df	P-value
	7.7827	5	0.1686

Table 10: Granger test with Regulations (REG) as dependent variable, the test shows that REG does Granger cause the other variables but that there is not an instant effect

Regulations (REG) does Granger cause the other variables which makes sense, for instance, if the limit for approved housing cost increases, it will become easier to get housing allowance. No instantaneous causality is detected.

Dependent variable AGE

$Granger	F-test	df1	df2	P-value
	1.5159	15	390	0.09618

$Instant	Chi-squared	df	P-value
	21.377	5	0.0006873

Table 11: Granger test with Age (AGE) as dependent variable, the test shows that AGE does not Granger cause the other variables but that there is an instant effect

Age does not Granger cause the other variables. However, we can detect instantaneous causality. This might be a result of the declining average age we detected in the descriptive part of the analysis.

Dependent variable EMP

$Granger	F-test	df1	df2	P-value
	2.2528	15	390	0.004782

$Instant	Chi-squared	df	P-value
	13.379	5	0.02007

Table 12: Granger test with Employment (EMP) as dependent variable, the test shows that EMP does Granger cause the other variables and that there is an instant effect

Employment (EMP) does Granger cause the other variables. Meaning that the level of employment in Norway helps predict the number of new recipients. Moreover, it has an instantaneous effect.

Dependent variable UMP

$Granger	F-test	df1	df2	P-value
	1.941	15	390	0.01841

$Instant	Chi-squared	df	P-value
	9.921	5	0.0775

Table 13: Granger test with Unemployment (UMP) as dependent variable, the test shows that UMP does Granger cause the other variables but that there is not an instant effect

As with employment (EMP), unemployment (UMP) Granger causes the other variables and has an instantaneous effect on the level of new recipients (NEW).

6.6.4 Variable Impact Analysis

The first method we are going to use is the Impulse response Function (IRF). This method analyses the response to a unit shock of another variable. We are going to look at the effect that regulation, housing expenses, unemployment and employment has on the number of new recipients. The generalized IR model are supposed to “replicate” the effect that a given shocks has, simulating and plotting the effect in the graph. The IRF model is based on the VAR model that we have created.

The second method we are going to use is the forecast error variance decomposition (FEVD). The FEVD tables are normalized values of the variables used in the VAR model. Which enables us to determine how much of the variance on the dependent variable is lagged by its own variance. Additionally, it shows how much the other variables in the model explains the variability in the chosen variable over time. The FEVD is an extension of the IRF method. Since we know that not all of the variance is explained by a single variable, we look at the FEVD to see the effect that each variable has on the other variables. We will use a timeframe of 10 months.

The first simulation is a positive shock from the regulation variable (REG). Meaning that the number of regulations increases. The results show how this effect the inflow of new recipients. The simulation in the figure below shows that an increased number of regulations leads to an immediate increase in number of new recipients. After the immediate increase, the inflow of new recipients significantly drops after three months before it peaks around month 7. After this it drops below zero before it normalises.

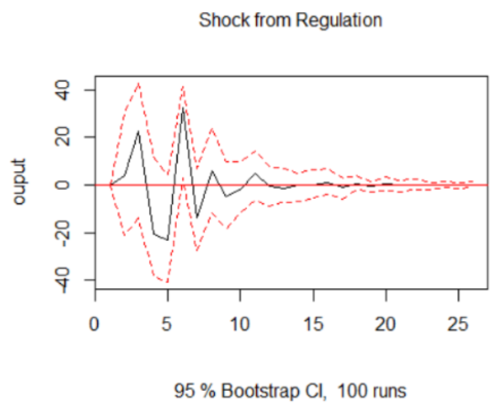


Figure 33: Impact simulation of a positive shock from Regulation (REG)

It must be mentioned that the confidence interval for the positive shock of regulations is quite large in the first months, indicating that the actual direction of the development of new recipients is not entirely sure. However, as we remember from the descriptive part of this study, most regulations tend to be an increase in the maximum limit of approved household expenses and an increase in the maximum income limit which leads to an instant effect, before it flattens out.

The housing expenses shock, shown in the figure below, involves a much shorter period of influence on new receivers compared to regulation. The simulation tells us that an increase in housing expenses leads to a lower number of new recipients. Only after 5 months are the effects of the shock normalized, we experience some variation, but it is minimal. However, the confidence interval has a large gap up till 10 months after the shock. Which indicate that there are uncertainty concerning the estimate, the response could be both positive and negative.

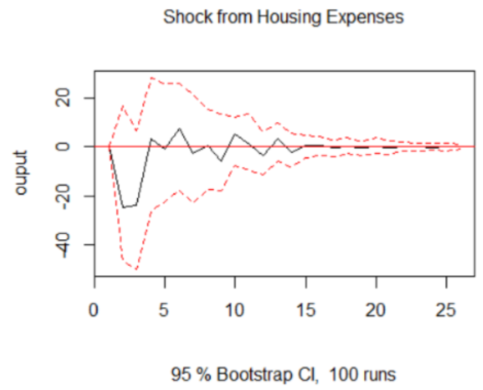


Figure 34: Impact simulation of a positive shock from housing expenses (AHE)

Moreover, we simulate a positive shock of unemployment (UMP). In our case, a positive shock, means that the unemployment rate increases. The simulation shows that an increase in unemployment decreases the number of new recipients. This is contradictory to what is expected; higher unemployment leads to an increase in number of recipients. However, there is a large gap between the predicted response and the confidence interval, which indicate that there is great uncertainty associated with the predicted response.

Figure 35: Impact simulation of a positive shock from unemployment (UMP)

In the last simulation, we illustrate a positive shock from employment. Which means that more people are employed, in our case, this leads to a certain decrease in the number of new recipients. However, after only 2 to 3 months the number of new recipients starts increasing again, eventually stabilizing.

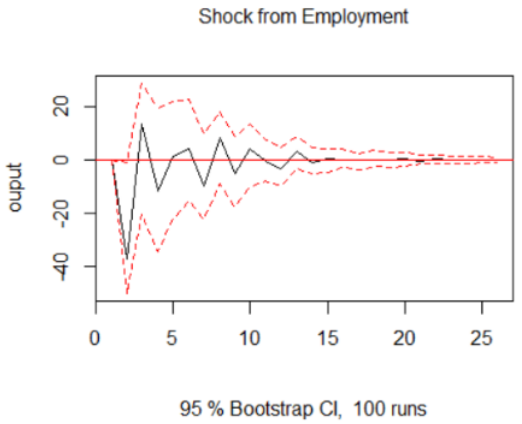


Figure 36: Impact simulation of a positive shock from employment (EMP)

6.6.5 Forecast error variance decomposition

The most interesting variable for us to look closer into is the New recipients, since we are forecasting new recipients. If we take closer look into FEVD of new recipients, we can see that the response of new recipients is mainly caused by themselves. However, we can see after the first-time horizon, the other variables start to play a little part of the variance decomposition. Where employment is the variable with the highest effect in the first 5 periods (5%). This support the findings in the IRL analysis, where employment has the most precise estimate of a shock. However, the importance of regulations variables increases to eventually be the main force driving the new recipients upwards. The main result from the FEVD analysis is that employment is the forcing driver for the variance in new recipients the effect stays almost constant through a time period of 10 months, where a shock in regulation will have a greater impact, with a lagged effect.

Dependent variable NEW

Month	NEW	AHE	AGE	REG	UMP	EMP
[1,]	1.0000000	0.00000000	0.000000000	0.000000000	0.00000000	0.00000000
[2,]	0.8883042	0.02368908	0.008657399	0.000659166	0.02632924	0.05236088
[3,]	0.8532535	0.04035301	0.011457283	0.017859427	0.02485108	0.05222574
[4,]	0.8305445	0.03864790	0.020624973	0.030923973	0.02520485	0.05405380
[5,]	0.8091942	0.03749453	0.028299508	0.046779762	0.02574350	0.05248852
[6,]	0.7793221	0.03787463	0.029624246	0.076883396	0.02527758	0.05101807
[7,]	0.7725440	0.03776079	0.029343649	0.081709200	0.02516483	0.05347753
[8,]	0.7693132	0.03760608	0.030184081	0.082458729	0.02509913	0.05533876
[9,]	0.7652296	0.03846127	0.032415214	0.082765676	0.02532619	0.05580208
[10,]	0.7620637	0.03907552	0.034949298	0.082481419	0.02537611	0.05605394

Table 14: Forecast error variance decomposition for variable new recipients (NEW)

6.6.6 Model performance

To measure the model performance, we have split the data into a training- and a test set. Where we have excluded the last twelve months in the training data and use these twelve months as a test set, so we could compare the predicted forecast up against the test values. In addition, have we included a simple linear regression with season- and trend components, to see how this basic linear regression model performs compared to our VAR model. We have also calculated the mean error (ME), root mean squared error (RMSE), mean absolute error (MAE), mean percentage error (MAE) and Mean absolute percentage error (MAPE) for the VAR model and the linear regression model.

As we can see from figure 37, both models do well in predicting the number of new recipients the next two years. The first we can see, is that the linear regression constantly overestimate the number of new recipients (besides of august 2019 and January 2020). Whereas the VAR model underestimate mostly of the months. We can see that both the Linear regression and the Vector Autoregression follow a similar pattern, the reason for this is that both models includes a trend and seasonal components. However, we can clearly see that our VAR model does a better job in general to forecast the number of new recipients.

Date	Linear Regression	VAR	Actual
2019.04	1723	1496	1328
2019.05	1453	1231	1345
2019.06	1441	1256	1172
2019.07	1310	1050	1201
2019.08	1347	1128	1445
2019.09	1685	1450	1501
2019.10	1708	1478	1582
2019.11	1783	1558	1534
2019.12	1463	1207	1055
2020.01	1647	1358	1651
2020.02	1848	1488	1391
2020.03	1710	1423	1403

Table 15: Forecasted values compared to actual values of the test set. The year and months are on the left, whereas the forecasted values with VAR, Linear Regression, and actual values are on the right. The Linear regression constantly overestimates, whereas the VAR model underestimate most of the months.

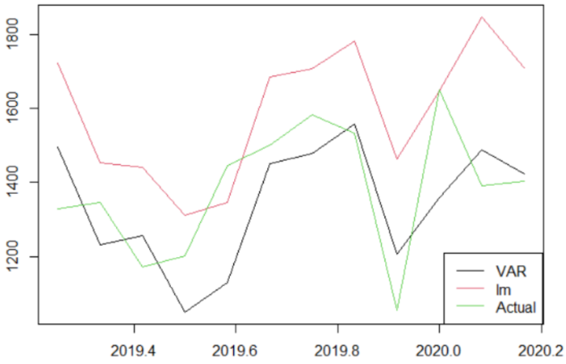


Figure 37: Forecasted values compared to actual values. The green line is the actual values, the black line is the forecasted values with VAR, and the red line is the simple linear regression. The Linear regression constantly overestimates, whereas the VAR model underestimate most of the months.

All the accuracy measures of the forecasts are shown in the table 18. The VAR model has a MAE of 131, whereas the linear regression has 226. This means that the VAR model on average estimate 131 new receivers wrong, while the linear regression estimates on average 226 wrong. The RMSE is 159 for the VAR model and 265 for the linear regression. The last measurement we will go closer into is the MAPE, where the var model has 9 and the linear regression has 17. Which means that the VAR model have an error of approximately 10%, whereas the linear regression has almost twice as high error with 17% When we compare both models against the actual values. This support what we could see in figure 37, that the VAR model on average does a better job in predicting the future number of recipients.

Accuracy Measurement	Vector autoregression	Linear regression
ME	39,88	-209,16
RMSE	159,11	265,02
MAE	131,23	226,17
MPE	2,17	-16,05
MAPE	9,57	17,22

Table 16: Accuracy results of the test set. The accuracy measurements are on the left, and the results for the VAR model and Linear Regression model are on the right. As the figure displays the VAR model consistently beats the simple Linear Regression model in terms of accuracy.

6.6.6 Forecast April 2020 – December 2021

In our forecast we have used the same VAR model, with the same specifications as we used above. The only difference is that we include the test set, which means that we now have 111 observations to train our model on. When we ran the same tests as above, we could detect that there was tendency of serial correlation in the residual plot, which means that an error associated with a given period could carry over to the next period. However, the serial correlation was not that high. Hence, we decided to run the model with the same specifications. Subsequently, we must be careful when we interpret the result from the forecast since the tendency of serial correlation could lead to a forecast error.

The time-horizon for our forecast is 21 months, which means that we forecast the number of new recipients up till December 2021. The raw output of the historical data and the predicted forecast is shown in the figure below, the Y-axis represent the transformed values with a blue dotted line (the first differences) and the 95% confidence interval is represented by the red dotted lines. Both historical- and predicted forecast values have a high volatility, which makes it harder to get an accurate forecast. Additionally, we can detect that the confidence interval starts to get bigger/wider after approximately 10 months, which indicate that there is uncertainty around the forecast after that point. The predicted forecast is consistent with the historical values of new recipients. However, we can see that the forecast does not reach the same level as the positive peaks, but the 95% confidence interval does, which lead it to a possibility that the number of new recipients could reach that levels. To get a better overview of the actual number of new recipients, we will transform the values back to their original form, which makes it more intuitively to interpret.

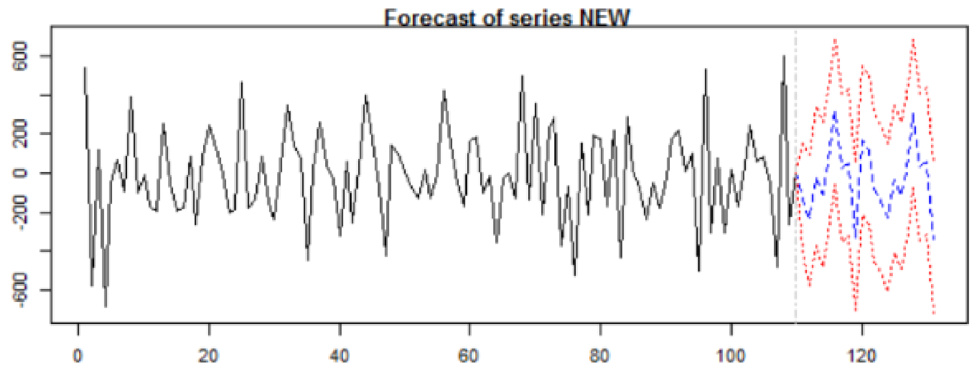


Figure 38: The raw output of the historical data and the predicted forecast is shown in the figure below, the Y-axis represent the transformed values with a blue dotted line (the first differences) and the 95% confidence interval is represented by the red dotted lines

Below we can see the monthly numbers of new recipients, in its original form. A significant takeaway from figure 39, is that we have a drastic drop the first few months in our forecast, this lasts for approximately 4 months. After that there is a significant growth the next few months. After roughly 10 months we can see a dramatic drop, where the monthly number of new recipients is as low as 900. This is an all time low for new recipients. We also must consider the serial correlation tendency in our model. This could lead to a lower number of new recipients in our forecast compared to the actual number. To get a better overview of the forecast in conjunction with the development of the number of new recipients. The next table will include the number of new recipients since January. 2011 and the forecast.

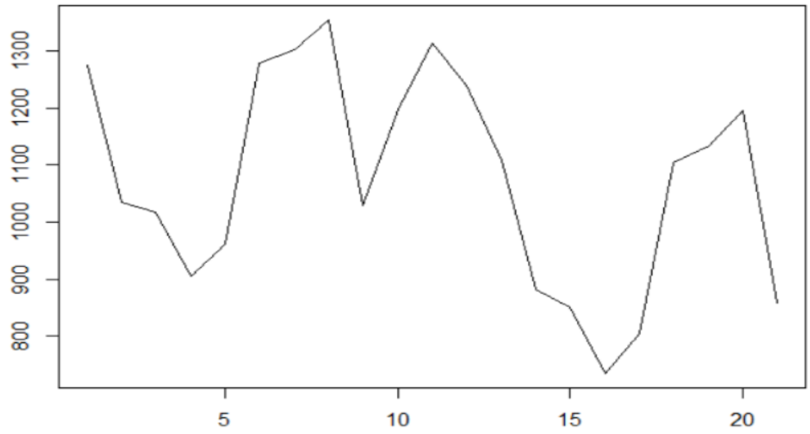


Figure 39: Monthly numbers of new recipients, in its original form. A significant takeaway from the output is that there is a drastic drop the first few months of the forecast, this persists for approximately 4 months. After that there is a significant growth the next few months.

In the figure below all the observations since 2011 are included plus the predicted forecast. There has been a downward trend since the change in income basis in 2017. As we can see our model tells us that this trend will continue. The reason for this is that in the past three years, the trend has approximately been the same. Consequently, this leads to strong incentives that the trend will keep on going the next 2 years. Furthermore, it must be taken into consideration that our model cannot predict changes based on future changes, meaning that if the NSHB makes significant changes in their regulations regarding housing allowance as a result of the decreasing trend it will not show in our model.

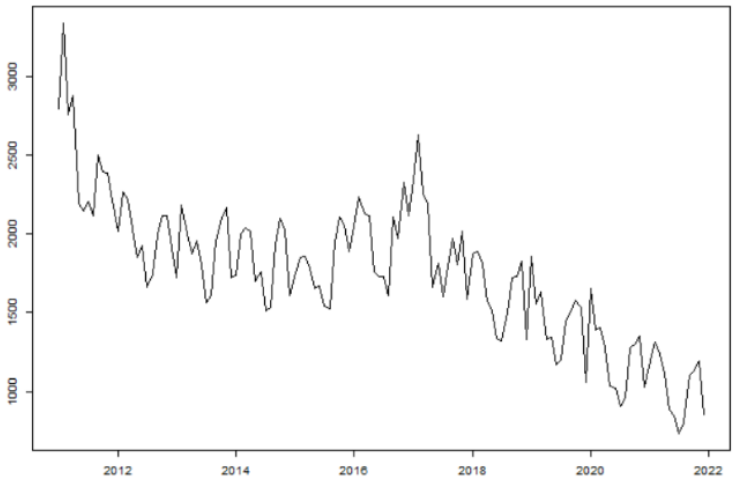


Figure 40: Actual observations January 2010 - March 2020 and predicted values from April 2020 - 31.12.2021. There has been a downward trend since the change in income basis in 2017. As we can see, our model tells us that this trend will continue.

There are many factors that could potentially affect the future number of new recipients. Most of the macroeconomic variables are given on a yearly or quartile level. This have limited our use of external data in the forecast. A yearly forecast was considered; however, a yearly forecast would only have provided us with ten observations per time series and was therefore considered not sufficient. A quarterly forecast was also considered, but in the end, we decided to forecast on a monthly basis. Thus, providing us with 111 observations per time series.

Furthermore, data including the ongoing pandemic with the corona virus, is not part of this forecast. The pandemic has had a significant effect on the Norwegian population, labour market, and the disadvantaged. It is likely that the pandemic will affect our prediction, therefore data from this period should be added in the future to obtain a better forecast.

7. Conclusion

In this study we analysed the development of housing allowance recipients in Norway for the past ten years. The main focus has been on new recipients, how they have developed, what affects the inflow, both in terms of internal and external factors, and how the development will be in the near future. To answer these questions, we have performed a descriptive analysis of applicants, recipients, and new recipients. This was done to see how new recipients differ from applicants and recipients in general. To determine which variables, have the biggest effect on the inflow, impulse response function (IRF) and FEVD was used. Lastly, a statistical model, capable of predicting the inflow of new housing allowance recipients was developed by using the econometric technique, Vector Autoregression.

Based on our descriptive analysis of applicants and housing allowance recipients, we find there to be a clear downward trend the past 7 years. Especially in municipality group 4. The decline from 2017 to 2020 can partly be explained by the change in income basis calculation which lead to a higher rate of declines. Additionally, the external data shows that the number of people on Work clarification benefit has decreased by 30% from 2012 – 2020. The number of people with disability benefits considered poor was just 12% and the number of pensioners considered poor has decrease significantly since 2008. The study confirms the assumption made by Fjelltoft & Ezat (2019) that new recipients of housing allowance to a greater extent will be people with very low or no income, this is especially true for people living in municipality group 1.

The research identified employment and regulation as the key determinants of the inflow of new recipients. However, the development concerning the number of new recipients is mainly caused by itself. The evidence from the predicted forecast suggest that the VAR model tracks the movement in actual values fairly well. Moreover, the forecast suggests that there will be a lower number of new recipients the next 21 months. However, there are uncertainty around the forecast which could indicate that there are other factors that are not included in the model that could impact the outcome. In addition to the uncertainty, our data does not include data from the corona pandemic. So far, superficial analysis conducted by the NSHB has not showed huge consequences from the corona pandemic, but one risks that our conclusion might be redundant. Consequently, future research is needed to continue the investigation on what affects the inflow of new recipients and data including numbers from the corona pandemic should be included to see how the model fits under the changed circumstances.

8. Bibliography

Akashi, F., Odashima, H., Taniguchi, M. et al. A New Look at Portmanteau Tests. Sankhya A 80, 121–137 (2018). Retrieved from: https://doi.org/10.1007/s13171-017-0109-3

Amundsen, B. (2019, October 31). Litt flere fattige i Norge. Retrieved from https://forskning.no/penger-ulikhet-okonomi/litt-flere-fattige-i-norge/1585230?fbclid=IwAR17icFK9ztQ3Tt9cSTY8UZb2ZFeaWdtiFsJjLq7KxdEhJAskuwpJzlGYrI

Arnold, M., Hanck, C., Gerber, A., & Schmelzer, M. (2020). Introduction to Econometrics with R. Retrieved from https://www.econometrics-with-r.org/index.html

Brooks, C., & Tsolacos, S. (2010). Real estate forecasting in practice. *Real Estate Modelling and Forecasting,* 414-433. doi:10.1017/cbo9780511814235.014

Chai, T., & Draxler, R. R. (2014). Root mean square error (RMSE) or mean absolute error (MAE)? – Arguments against avoiding RMSE in the literature. *Geoscientific Model Development, 7*(3), 1247-1250. doi:10.5194/gmd-7-1247-2014

Chambers, J. C., Mullick, S. K., & Smith, D. D. (1971). How to Choose the Right Forecasting Technique. Retrieved from https://hbr.org/1971/07/how-to-choose-the-right-forecasting-technique

Christensen, J., & Berg, O. (2019, September 04). Velferdsstat. Retrieved from https://snl.no/velferdsstat

Diebold, F., & Lopez, J. (1996). Forecast Evaluation and Combination. doi:10.3386/t0192

Fjelltoft, A., & Ezat, H. (2019). Dagens bostøttemottakere, En gjennomgang av bostøttemottakere i 2019, sett i lys av endringer fra 2009 - 2019. Retrieved from https://www.husbanken.no/rapporter/?pg=36

Formplus. (2020, March 12). What is Empirical Research Study? [Examples & Method]. Retrieved from https://www.formpl.us/blog/empirical-research

Glen, S. (2016, June 07). "ADF — Augmented Dickey Fuller Test". Retrieved from: https://www.statisticshowto.com/adf-augmented-dickey-fuller-test/

Glen, S. (2016, December 30). "Granger Causality: Definition, Running the Test". Retrieved from: https://www.statisticshowto.com/granger-causality/

Glen, S. (2016, Dexember 14). "Unit Root: Simple Definition, Unit Root Tests". Retrieved from: https://www.statisticshowto.com/unit-root/

Kaldheim, O. (n.d.). CV Osmund Kaldheim. Retrieved from https://www.regjeringen.no/globalassets/upload/kilde/krd/prm/2005/0394/ddd/pdfv/248773-cv_kaldheim.pdf

Kenton, W. (2020, April 26). What Is the Welfare State? Retrieved July 14, 2020, from https://www.investopedia.com/terms/w/welfare-state.asp

Kommunal- og moderniseringsdepartementet. (n.d.). Housing for welfare - National strategy for housing and support services (2014-2020). Retrieved from https://www.regjeringen.no/no/dokumenter/Housing-for-welfare---National-strategy-for-housing-and-support-services-2014-2020/id2351089/

Leamer, E. (1985) Vector Autoregressions for Causal Inference?, in K. Brunner – A.H. Meltzer (a cura di), Understanding Monetary Regime, Carnegie-Rochester Conference Series on Public Policy, 22, pp. 255-304.

Ministry of Local Government and Modernisation. (n.d.). Ministry of Local Government and Modernisation. Retrieved from https://www.regjeringen.no/en/dep/kmd/id504/

NAV. (2020, May 07). Fakta om NAV. Retrieved from https://www.nav.no/no/nav-og-samfunn/om-nav/fakta-om-nav

Norwegian State Housing Bank. (n.d.). Adequate and secure housing for all. Retrieved from https://www.husbanken.no/english/norwegian-housing-policy/

Norwegian State Housing Bank. (n.d.). Organization. Retrieved from https://www.husbanken.no/english/organisation/

Norwegian State Housing Bank. (n.d.). The History of the Norwegian State Housing Bank. Retrieved from https://www.husbanken.no/english/the-history-of-the-norwegian-state-housing-bank/

Norwegian State Housing Bank. (n.d.). The Norwegian State Housing Bank' role. Retrieved from https://www.husbanken.no/english/about-the-housing-bank/the-housing-banks-role/

NTB. (2019, December 16). Leieprisene i Oslo og Bærum klart over landsgjennomsnittet. Retrieved from https://www.nettavisen.no/okonomi/leieprisene-i-oslo-og-baerum-klart-over-landsgjennomsnittet/3423896521.html

Olje- og Energidepartementet. (n.d.). Norsk oljehistorie på 5 minutter. Retrieved from https://www.regjeringen.no/no/tema/energi/olje-og-gass/norsk-oljehistorie-pa-5-minutter/id440538/

Peixeiro, M. (2020, May 19). The Complete Guide to Time Series Analysis and Forecasting. Retrieved from: https://towardsdatascience.com/the-complete-guide-to-time-series-analysis-and-forecasting-70d476bfe775

Pfaff, B. (n.d.). Vars v1.5-3. Retrieved from: https://www.rdocumentation.org/packages/vars

Rouse, M., & Haughn, M. (2017, August 07). What is empirical analysis? - Definition from WhatIs.com. Retrieved from https://whatis.techtarget.com/definition/empirical-analysis

Singh, A. (2020, May 07). Multivariate Time Series: Vector Auto Regression (VAR). Retrieved from https://www.analyticsvidhya.com/blog/2018/09/multivariate-time-series-guide-forecasting-modeling-python-codes/?utm_source=DataCamp.com

Skiphamn, S. (2020, February 27). Innsikt: Lever enslige minstepensjonister under fattigdomsgrensen? Retrieved from: https://www.faktisk.no/artikler/J06/innsikt-lever-enslige-minstepensjonister-under-fattigdomsgrensen

SSB. (2018, October 17). Dette bruker nordmenn penger på. Retrieved from https://www.ssb.no/nasjonalregnskap-og-konjunkturer/artikler-og-publikasjoner/dette-bruker-nordmenn-penger-pa

SSB. (2019). *Får de som har rett på det bostøtte?* (Rep. No. 2019/2). Statistisk Sentralbyrå.

SSB. (2019, December 16). Income and wealth statistics for households. Retrieved from https://www.ssb.no/en/inntekt-og-forbruk/statistikker/ifhus/aar

SSB. (2020, July 23). Labour force survey, seasonally-adjusted figures. Retrieved from https://www.ssb.no/en/arbeid-og-lonn/statistikker/akumnd

SSB. (2020, July 02). Recipients of disability benefit. Retrieved from: https://www.ssb.no/sosiale-forhold-og-kriminalitet/statistikker/ufore/aar

SSB. (2019, December 16). Rental market Survey. Retrieved from: https://www.ssb.no/en/priser-og-prisindekser/statistikker/lmu

SSB. (n.d.). Some remarks on the history of official statistics in Norway. Retrieved from: https://www.ssb.no/en/omssb/om-oss/historie

Stock, J. H., & Watson, M. W. (2015). *Introduction to Econometrics, Third Update, Global Edition*. Pearson Education Limited.

Willmott, C., & Matsuura, K. (2005). Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance. *Climate Research, 30*, 79-82. doi:10.3354/cr030079

World Bank Group. (2020). GDP per capita (current US$). Retrieved from: https://data.worldbank.org/indicator/NY.GDP.PCAP.CD?end=2019

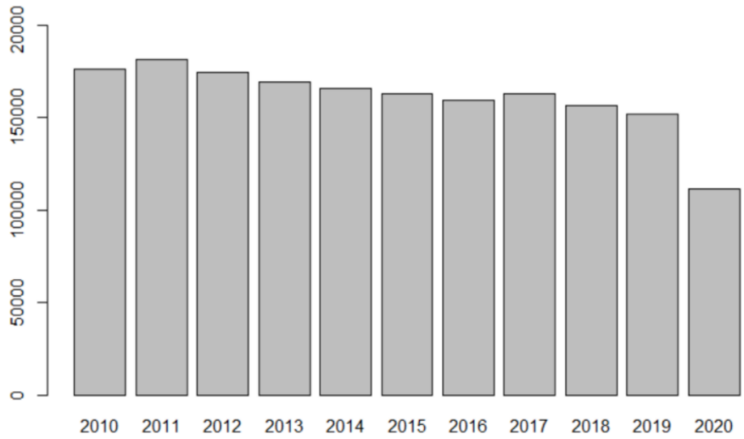
9. Appendix

Appendix – Table of figures

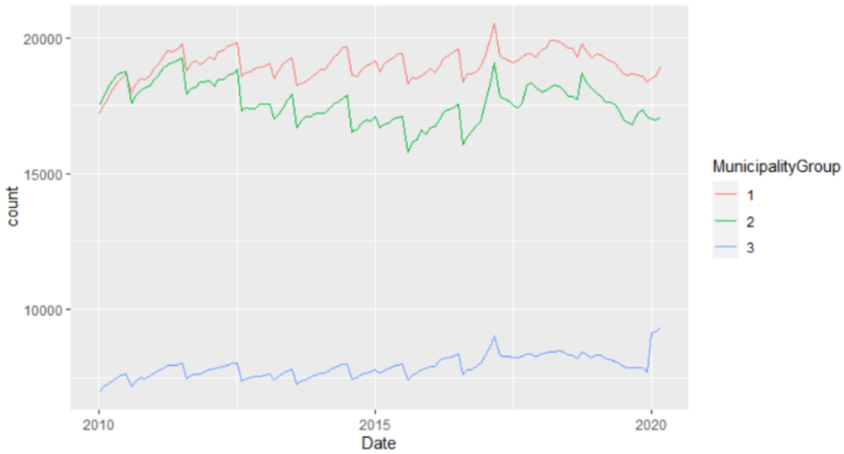
Appendix - Table of tables

A. Plots

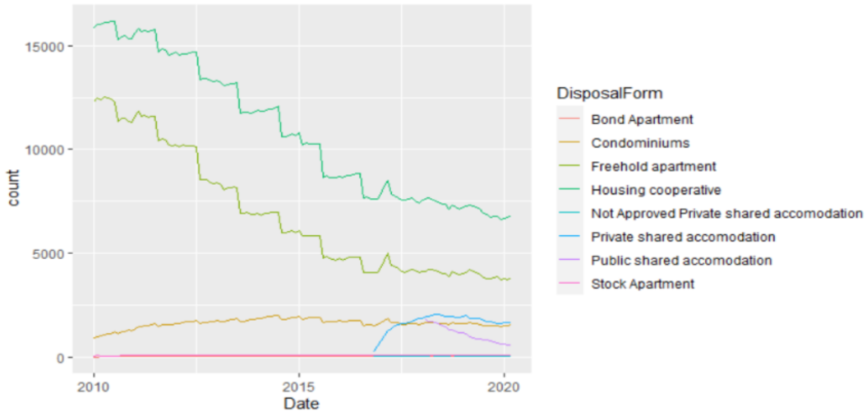
A.a Total number of received applications in the period 2010-2020



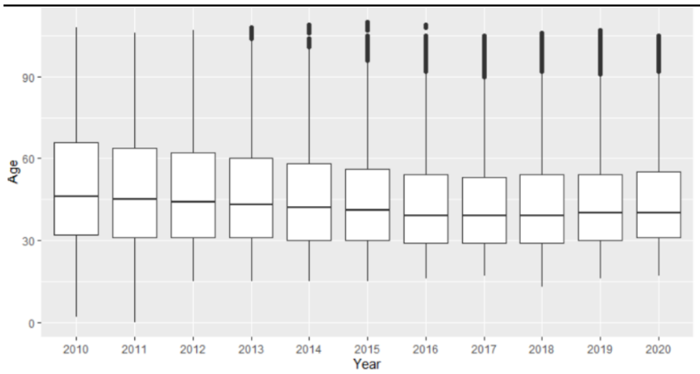
A-Figure 1: Aggregated yearly unique applications, 2011-2020. The Y-axis shows the number of unique applications whereas the X-axis shows years. Note that 2020 is not a complete year and only contains data from January 2020 – March 2020.



A-Figure 2: Development in total number of applications by municipality group 1-3, 2011-2020. The Y-axis shows the total number of unique applications per municipality group. The X-axis shows the years.

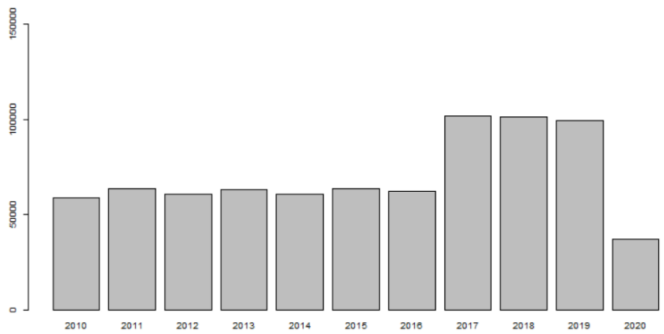


A-Figure 3: Development of applicants based on disposal form excluded privately rented housing and public housing, 2010-2020.

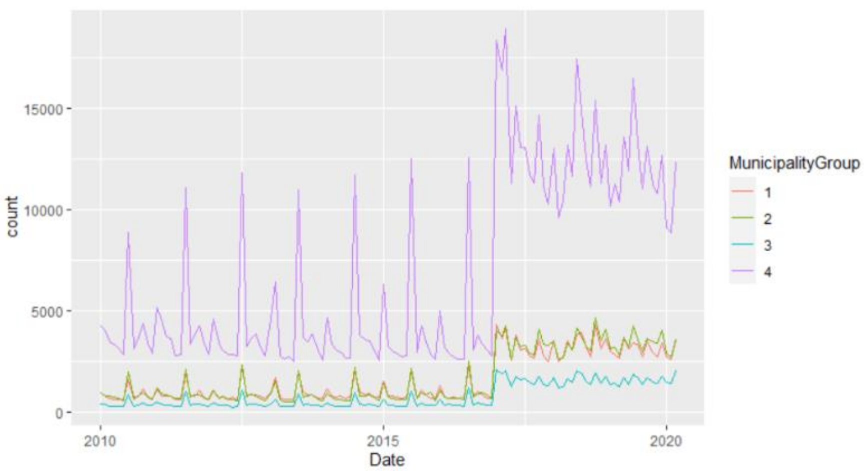


A-Figure 4: Applicant age yearly development from 2010-2020. The Y-axis represents the age and the X-axis shows years

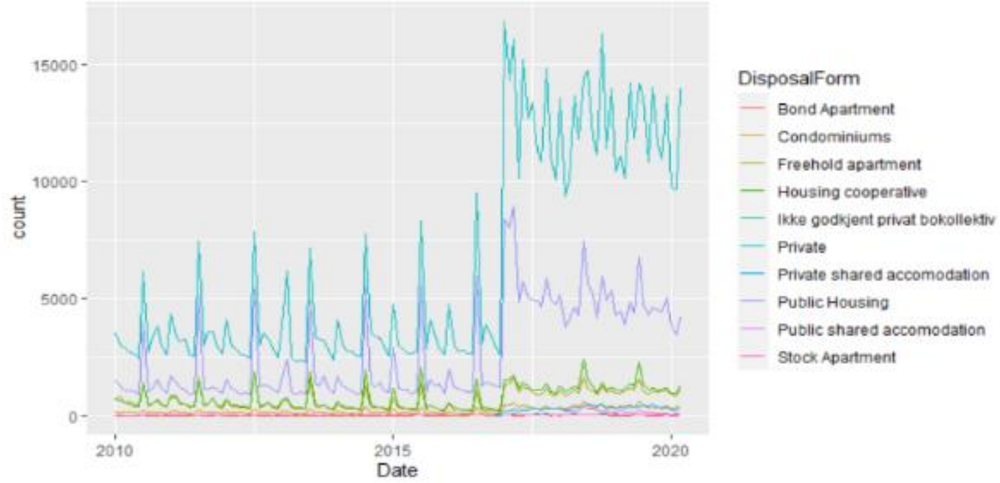
A.b Total number of declined housing allowance applications 2010-2020



A-Figure 5: Yearly housing allowance application rejections, 2010-2020. The Y-axis shows the total number of unique declines. The X-axis shows the years. Note that 2020 only contains data from January 2020 – March 2020 and is therefore not comparable with the other years.

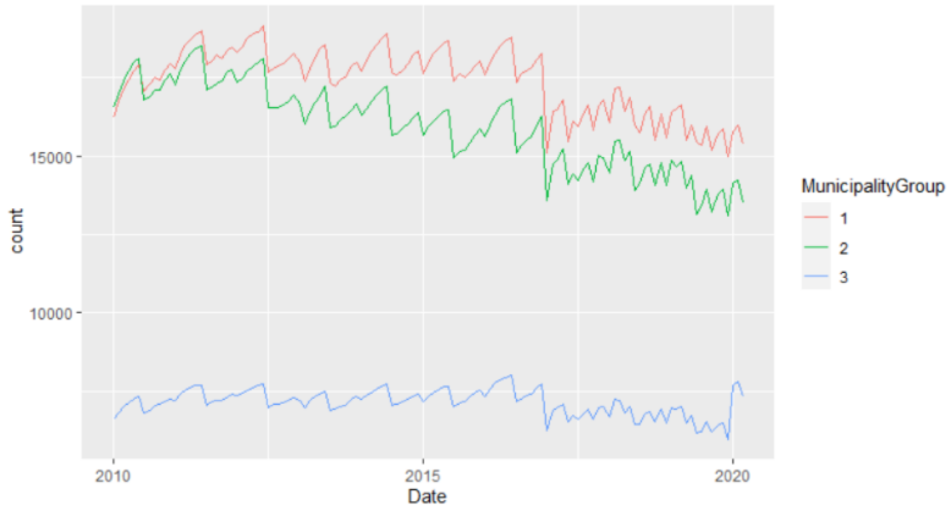


A-Figure 6: Application rejections by municipality group from 2010-2020. The Y-axis represent total number of rejections per group per month. The X-axis represents months.

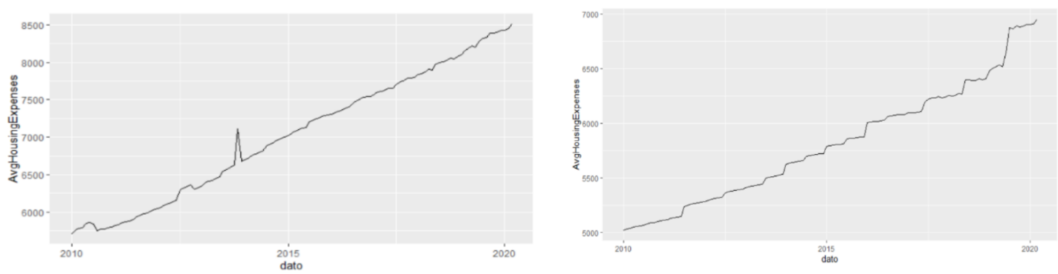


A-Figure 7: Total number of rejections based on disposal form from 2010-2020. The Y-axis represents the total number of rejections per group per month. The X-axis shows months.

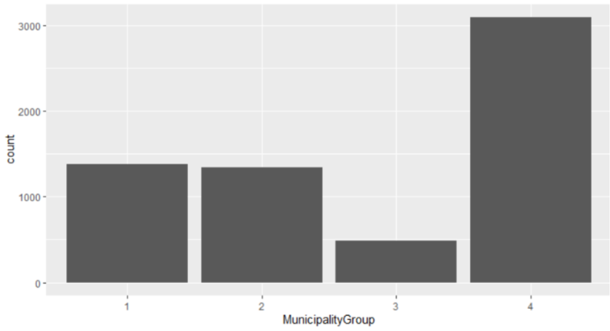
A.c Total number of recipients of housing allowance 2010-2020



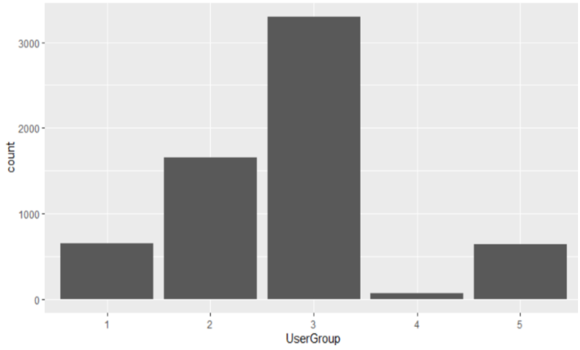
A-Figure 9: Monthly development of recipients by Municipality group 1-3, 2010-2020. The Y-axis shows the total number of recipients. The X-axis shows months



A-Figure 8: Development of monthly actual Housing Expenses (Left) vs development of monthly approved housing expenses (right), 2010-2020

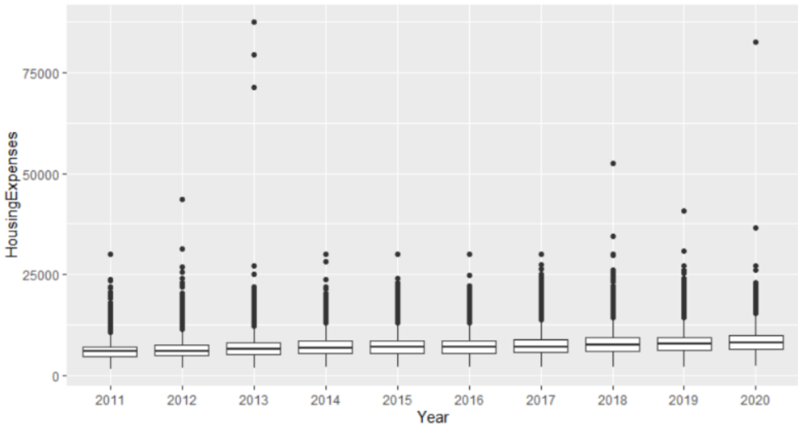


A-Figure 10: Distribution of recipients who have received housing allowance constantly from January 2010 - March 2020



A-Figure 11: Distribution of recipients who have received housing allowance constantly from January 2010 - March 2020

A.d New recipients of housing allowance 2011-2020



A-Figure 12: Average housing Expenses of new recipients by year, 2011-2020

B. Internal Data - descriptive statistics

Statistic	N	Mean	St. Dev.	Min	Pctl(25)	Pctl(75)	Max
VedtakSk	14,334,884	72,820,881.000	89,967,214.000	9,296,796	13,471,423.0	106,959,154.0	298,274,700
DatoIdTermin	14,334,884	20,144,868.000	29,479.550	20,100,101	20,120,401	20,170,601	20,200,301
HusstandSammensetningSk	14,334,884	1.635	1.042	1	1	3	4
BrukergruppeTrygdSk	14,334,884	3.574	1.349	1	3	5	5
AntallHusstandsmedlemmerSk	14,334,884	1.592	1.091	1	1	2	5
InntektIntervallId	14,334,884	1,000,010.000	792.387	0	1,000,008	1,000,012	1,000,946
AlderIntervallId	14,334,884	2.704	0.875	0	2	3	4
BeregnetBostotte	14,334,884	2,125.563	1,265.841	0	1,221	2,880	7,733
BeregnetBoutgift	14,334,884	5,773.916	1,312.212	0	5,083	6,625	12,236
SamletInntekt	14,334,884	152,723.900	137,825.400	0	95,622	194,208	97,246,338
HusstandId	14,334,884	638,207.600	365,604.400	3	328,992	946,356	1,357,469
AntallHusstandsmedlemmer	14,334,884	1.629	1.237	1	1	2	16
AntallHusstandsmedlemmer10	14,334,884	1.629	1.235	1	1	2	10
OvergangsordningSk	14,334,884	0.200	0.502	0	0	0	2
KommuneSk	14,334,884	1,036.254	1,293.282	2	156	1,270	6,687
DisposisjonsFormSk	14,334,884	4.327	1.944	3	3	4	17
ErElektroniskSoknad	14,334,884	0.020	0.141	0	0	0	1
Alder	14,334,884	45.891	19.355	0	30	58	110
SumBoutgift	14,334,884	6,981.531	15,771.580	0	5,295	8,111	50,979,622
RimeligBoutgift	14,334,884	5,140.347	154,037.400	1,250	1,691	3,736	203,625,436
Kommunegruppe	14,334,884	3.140	1.184	1	2	4	4
TerminÃ.r	14,334,884	2,014.423	2.950	2,010	2,012	2,017	2,020
TerminMÃ.ned	14,334,884	6.348	3.468	1	3	9	12
FormueTillegg	14,334,884	2,087.156	57,332.660	0	0	0	81,849,925
KommuneNr	14,334,884	958.684	836.194	101	319	1,247	5,444
BelopLysOgVarme	5,063,150	503.177	7.539	500.000	500.000	500.000	528.000
AntallKvadratmeter	639,312	63.293	26.900	0.000	49.000	72.000	880.000
AntallRomTotalt	165,520	2.482	1.152	0.000	2.000	3.000	48.000

A-Table 1: Descriptive statistics on internal data

C. Rental Market Survey

	Average annual rents per sqm							
	2012	2013	2014	2015	2016	2017	2018	2019
The whole country								
1 room	2 310	2 510	2 610	2 660	2 870	2 920	3 040	3 100
2 rooms	1 550	1 670	1 750	1 830	2 050	2 120	2 180	2 280
3 rooms	1 310	1 440	1 490	1 570	1 750	1 820	1 900	1 960
4 rooms	1 140	1 220	1 230	1 320	1 530	1 610	1 710	1 740
5 rooms or more	940	980	1 040	1 050	1 270	1 450	1 500	1 550

A-Table 2: Average rents are weighted and can therefore not be added up. Number of observations differ according to stratification. Average rent levels between years are not directly comparable since the survey is based on unique samples each year that can differ according to variables that are important for the rent level (SSB, 2019).

D. Income and Wealth statistic

	Total income, median (NOK)								
	2010	2011	2012	2013	2014	2015	2016	2017	2018
All households	528 000	550 000	570 000	593 000	610 000	628 000	635 000	648 000	666 000

A-Table 3: 'All households' includes observations which are not included in the household types 'living alone', 'couple without resident children', 'couple with resident children 0-17 year' and 'single mother/father with children 0-17 year'. Couples include married couples, cohabiting couples and registered partners. Student households and children below the age of 18 years who are living alone, are excluded (SSB, 2018).

E. Recipients of Disability benefit

	Recipients of disability benefit								
	2012	**2013**	**2014**	**2015**	**2016**	**2017**	**2018**	**2019**	**2020**
Both genders	310967	310855	317136	321300	329371	321300	329371	343438	355224

A-Table 4: Recipients of disability benefit, the statistic includes persons who are registered with a positive decision on disability benefit during a calendar year. For the majority of the statistics, there is also a condition of being registered as a resident at the end of the year.

F. Recipients of Work Clarification Benefit

	Work clarification benefit (AAP)								
	2012	**2013**	**2014**	**2015**	**2016**	**2017**	**2018**	**2019**	**2020**
Both genders	170249	167815	152538	152537	148428	146786	138480	122757	118052

A-Table 5: Recipients of work clarification benefit. The numbers in the table represents recipients in March each year.

G. Employment

A-Table 6: Employment, seasonally adjusted, 3-months moving average by contents and month (SSB, 2020). The original table was in units of 1000, we have altered this, so it shows unit = 1.

Employed persons					
Year	Month	Value	Year	Month	Value
2011	1	2513000	2015	9	2641000
2011	2	2520000	2015	10	2638000
2011	3	2526000	2015	11	2634000
2011	4	2533000	2015	12	2645000
2011	5	2534000	2016	1	2648000
2011	6	2540000	2016	2	2645000
2011	7	2541000	2016	3	2640000
2011	8	2554000	2016	4	2637000
2011	9	2552000	2016	5	2635000
2011	10	2554000	2016	6	2638000
2011	11	2558000	2016	7	2641000
2011	12	2564000	2016	8	2643000
2012	1	2569000	2016	9	2649000
2012	2	2574000	2016	10	2643000
2012	3	2579000	2016	11	2632000
2012	4	2586000	2016	12	2634000
2012	5	2581000	2017	1	2632000
2012	6	2577000	2017	2	2640000
2012	7	2576000	2017	3	2639000
2012	8	2587000	2017	4	2641000
2012	9	2594000	2017	5	2645000
2012	10	2594000	2017	6	2638000
2012	11	2590000	2017	7	2644000
2012	12	2588000	2017	8	2646000
2013	1	2588000	2017	9	2655000
2013	2	2593000	2017	10	2659000
2013	3	2596000	2017	11	2652000
2013	4	2594000	2017	12	2655000
2013	5	2592000	2018	1	2663000
2013	6	2601000	2018	2	2679000
2013	7	2609000	2018	3	2680000
2013	8	2613000	2018	4	2681000
2013	9	2610000	2018	5	2689000
2013	10	2618000	2018	6	2697000
2013	11	2613000	2018	7	2700000
2013	12	2609000	2018	8	2702000
2014	1	2609000	2018	9	2704000
2014	2	2612000	2018	10	2710000
2014	3	2626000	2018	11	2708000
2014	4	2626000	2018	12	2704000
2014	5	2632000	2019	1	2703000
2014	6	2621000	2019	2	2712000
2014	7	2626000	2019	3	2719000
2014	8	2624000	2019	4	2718000
2014	9	2625000	2019	5	2712000
2014	10	2619000	2019	6	2724000
2014	11	2632000	2019	7	2730000
2014	12	2640000	2019	8	2744000
2015	1	2644000	2019	9	2737000
2015	2	2639000	2019	10	2731000
2015	3	2637000	2019	11	2733000
2015	4	2644000	2019	12	2736000
2015	5	2648000	2020	1	2749000
2015	6	2639000	2020	2	2747000
2015	7	2637000	2020	3	2733000
2015	8	2638000			

H. Unemployment

A-Table 7: Unemployment, seasonally adjusted, 3-months moving average by contents and month (SSB, 2020). The original table was in units of 1000, we have altered this, so it shows unit = 1.

Unemployed persons					
Year	Month	Value	Year	Month	Value
2011	1	89000	2015	9	127000
2011	2	83000	2015	10	131000
2011	3	87000	2015	11	132000
2011	4	89000	2015	12	142000
2011	5	90000	2016	1	138000
2011	6	89000	2016	2	138000
2011	7	91000	2016	3	134000
2011	8	89000	2016	4	133000
2011	9	93000	2016	5	133000
2011	10	91000	2016	6	132000
2011	11	93000	2016	7	133000
2011	12	92000	2016	8	131000
2012	1	93000	2016	9	129000
2012	2	86000	2016	10	123000
2012	3	82000	2016	11	116000
2012	4	82000	2016	12	113000
2012	5	86000	2017	1	118000
2012	6	84000	2017	2	123000
2012	7	85000	2017	3	126000
2012	8	90000	2017	4	125000
2012	9	92000	2017	5	120000
2012	10	96000	2017	6	115000
2012	11	95000	2017	7	112000
2012	12	98000	2017	8	110000
2013	1	99000	2017	9	109000
2013	2	102000	2017	10	111000
2013	3	102000	2017	11	111000
2013	4	101000	2017	12	111000
2013	5	98000	2018	1	108000
2013	6	102000	2018	2	110000
2013	7	103000	2018	3	107000
2013	8	103000	2018	4	109000
2013	9	100000	2018	5	108000
2013	10	104000	2018	6	111000
2013	11	108000	2018	7	112000
2013	12	101000	2018	8	111000
2014	1	96000	2018	9	111000
2014	2	92000	2018	10	103000
2014	3	93000	2018	11	103000
2014	4	90000	2018	12	107000
2014	5	95000	2019	1	107000
2014	6	97000	2019	2	101000
2014	7	104000	2019	3	93000
2014	8	102000	2019	4	95000
2014	9	106000	2019	5	102000
2014	10	104000	2019	6	108000
2014	11	107000	2019	7	106000
2014	12	114000	2019	8	110000
2015	1	119000	2019	9	108000
2015	2	121000	2019	10	112000
2015	3	119000	2019	11	110000
2015	4	123000	2019	12	106000
2015	5	129000	2020	1	101000
2015	6	127000	2020	2	101000
2015	7	126000	2020	3	101000
2015	8	124000			

I. R-script regarding the forecast

```
library(RODBC,lib.loc="c:/users/hqjb/library")
library(tidyverse,lib.loc="c:/users/hqjb/library")
library(crayon,lib.loc="c:/users/hqjb/library")
library(backports,lib.loc="c:/users/hqjb/library")
library(rstudioapi,lib.loc="c:/users/hqjb/library")
library(cli,lib.loc="c:/users/hqjb/library")
library(withr,lib.loc="c:/users/hqjb/library")
library(dplyr,lib.loc="c:/users/hqjb/library")
library(lubridate,lib.loc="c:/users/hqjb/library")
library(forcats,lib.loc="c:/users/hqjb/library")
library(dynlm,lib.loc="c:/users/hqjb/library")
library(zoo,lib.loc="c:/users/hqjb/library")
library(sandwich,lib.loc="c:/users/hqjb/library")
library(urca,lib.loc="c:/users/hqjb/library")
library(labeling,lib.loc="c:/users/hqjb/library")
library(farver,lib.loc="c:/users/hqjb/library")
library(digest,lib.loc="c:/users/hqjb/library")
library(vars,lib.loc="c:/users/hqjb/library")
library(tseries,lib.loc="c:/users/hqjb/library")
library(MTS,lib.loc="c:/users/hqjb/library")


# Connect to the database

con <- odbcDriverConnect('driver={SQL Server};server=sql-test02\\
atest;database=datavarehus;trusted_connection=true')

# Choose terms to process, term is p?? form YYYYMM01

for (TerminAr in c(2010,2011,2012,2013,2014,2015,2016,2017,2018,2
019,2020))

{

  for (TerminManed in c(01,02,03,04,05,06,07,08,09,10,11,12))

  {

    Termin <- (TerminAr * 10000) + (TerminManed * 100) + 1

    print(paste0("Behandler ??r:",toString(TerminAr),

                 " m??ned:",toString(TerminManed),

                 " termin:",toString(Termin)))


    query  <- paste0("

               SELECT   *

                       FROM [datavarehus].[bst].[VedtakPBI] v
```

```
                       where   v.DatoIdTermin between 20100101 an
d 20200301

    ")

    vt <- sqlQuery(con, query)

  }

}

odbcClose(con) # close connection




dim(vt)

str(vt)

count(vt)


theme_set(theme_classic())



#Remvove V00
vt1 <- vt[vt$Vedtakskode=="V00",]


##########################################
###### Forecasten ############
##########################################


#Formula to calculate new recipients

vt2 <- vt1%>%
  group_by(`HusstandId`)%>%
  mutate(date_of_first_engagement=min(`dato`))%>%
  ungroup()



vt2 <- vt2 %>%
  mutate(customer_Status = case_when(`dato`>date_of_first_engagem
ent~"Returning", `dato`==date_of_first_engagement~"New", TRUE ~"O
ther"))


new_and_returning_customers <- vt2 %>%
```

```
  group_by(floor_date(`dato`, unit = "month"))%>%
  summarise(New_Customers = n_distinct(`HusstandId`[customer_Stat
us=="New"]),

            Returning_customers = n_distinct(`HusstandId`[custome
r_Status=="Returning"] ), mean(`SamletInntekt`[customer_Status=="
New"]), mean(`BeregnetBoutgift`[customer_Status=="New"]), mean(`B
eregnetBoutgift`[customer_Status=="New"]), mean(`Alder`[customer_
Status=="New"]))
str(new_and_returning_customers)

new_and_returning_customers <- subset(new_and_returning_customers
, select = -c(Returning_customers ))
names(new_and_returning_customers)[1] <- "dato"
names(new_and_returning_customers)[2] <- "AntallMottakere"
names(new_and_returning_customers)[3] <- "Gjennomsnitlig Inntekt"
names(new_and_returning_customers)[4] <- "Gjennomsnitlig Boutgift
"
names(new_and_returning_customers)[5] <- "Gjennomsnittlig Alder"

new_and_returning_customers <- new_and_returning_customers[format
(new_and_returning_customers$dato,"%Y") !="2010", ]
new_receivers <- new_and_returning_customers

#External Data

reguleringer <- read.csv(file ="Reguleringer_a.csv",sep=";", head
er = TRUE, stringsAsFactors = TRUE)
reguleringer$?..Dato <- as.Date(reguleringer$?..Dato, format="%d.
%m.%Y")
names(reguleringer)[1] <-"dato"
reguleringer <- reguleringer[reguleringer[["dato"]]<="2020-03-01"
, ]

sysselsatte <- read.csv(file ="sysselsatte_m?nedlig...csv",sep=";
", header = TRUE, stringsAsFactors = TRUE)
sysselsatte$Year<- as.Date(sysselsatte$Year, format="%d.%m.%Y")
names(sysselsatte)[1] <- "dato"
sysselsatte <- subset(sysselsatte, select = -c(Category ))

arbeidsledige <- read.csv(file ="arbeidsledige_m?nedlig..csv",sep
=";", header = TRUE, stringsAsFactors = TRUE)

arbeidsledige$Year <- as.Date(arbeidsledige$Year, format= "%d.%m.
%Y")
names(arbeidsledige)[1] <- "dato"

arbeidsledige <- subset(arbeidsledige, select = -c(Category))


#Merging the external data with internal data

dplr <- left_join(new_receivers, reguleringer, by=c("dato"))
dplr <- left_join(dplr, arbeidsledige, by=c("dato"))
dplr <- left_join(dplr, sysselsatte, by=c("dato"))
```

```
dplr <- subset(dplr, select = -c(  Permitterte.regnet.som.arbeids
ledig.i.AKU,Permitterte.totalt.i.AKU,Sysselsatte.til.stede.p?.job
b.i.referanseuka,Utf?rte.ukeverk..a.37.5.timer.  ))
dplr <- subset(dplr, select = -c(`Gjennomsnitlig Inntekt`))
dplr <- subset(dplr, select = -c(dato))



names(dplr)[1] <-"NEW"
names(dplr)[2] <-"AHE"
names(dplr)[3] <-"AGE"
names(dplr)[4] <-"REG"
names(dplr)[5] <-"UMP"
names(dplr)[6] <-"EMP"


mymts = ts(dplr,
           frequency = 12,
           start = c(2011, 1))
mymts


trainingdata <- window(mymts, end=c(2019,3))
testdata <- window(mymts, start=c(2019,4))
plot(mymts)


# Main packages - problem: both have different functions VAR
## Testing for stationarity
### tseries - standard test adt.test
apply(mymts, 2, adf.test)




  stnry = diffM(trainingdata) #difference operation on a vector o
f time series. Default order of differencing is 1.

apply(stnry,2,adf.test)

plot.ts(stnry)

lagselect <- VARselect(stnry, lag.max = 10, type = "trend", seaso
n = 12)
lagselect$selection

autoplot(ts(stnry,
            start = c(2011,1),
            frequency = 12)) +
  ggtitle("Time Series Plot of the stationary `nye mottakere' Tim
e-Series")
```

```
# Lag order identification
#We will use two different functions, from two different packages
to identify the lag order for the VAR model. Both functions are q
uite similar to each other but differ in the output they produce.
vars::VAR is a more powerful and convenient function to identify
the correct lag order.
VARselect(stnry,
          type = "trend", #type of deterministic regressors to in
clude. We use none because the time series was made stationary us
ing differencing above.
          lag.max = 12,
          season= 12) #highest lag order


# Creating a VAR model with vars
var.a <- vars::VAR(stnry
                   ,
                   p = 3,
                   type= "trend",
                   season = 12

)

#
##TEsting the residuals
serial.test(var.a)

bv.arch <- arch.test(var.a, lags.multi = 12, multivariate.only =
TRUE)
bv.arch
bv.norm <- normality.test(var.a, multivariate.only = TRUE)
bv.norm
bv.cusum <- stability(var.a, type = "OLS-CUSUM")
plot(bv.cusum)

# a Shock, how does it affect New Recivers

irf.gdp <- irf(var.a, impulse = "REG", response = "NEW",
               n.ahead = 25, boot = TRUE)
plot(irf.gdp, ylab = "ouput", main = "Shock from Regulation")

irf.gdp1 <- irf(var.a, impulse = "AHE", response = "NEW",
                n.ahead = 25, boot = TRUE)
plot(irf.gdp1, ylab = "ouput", main = "Shock from Housing Expense
s")

irf.gdp2 <- irf(var.a, impulse = "UMP", response = "NEW",
                n.ahead = 25, boot = TRUE)
plot(irf.gdp2, ylab = "ouput", main = "Shock from Unemployment")

irf.gdp3 <- irf(var.a, impulse = "EMP", response = "NEW",
                n.ahead = 25, boot = TRUE)
plot(irf.gdp3, ylab = "ouput", main = "Shock from Employment")
```

```
irf.gdp <- irf(var.a, impulse = "Sysselsatte", response = "Antall
Mottakere",
               n.ahead = 25, boot = TRUE)
plot(irf.gdp, ylab = "ouput", main = "Shock from Arbeidsledige")

irf.gdp <- irf(var.a, impulse = "Sysselsatte", response = "Antall
Mottakere",
               n.ahead = 25, boot = TRUE)
plot(irf.gdp, ylab = "ouput", main = "Shock from Arbeidsledige")


#To generate the forecast error variance decompositions we make u
se of the fevd command, where we set the number of steps ahead to
ten.

bv.vardec <- fevd(var.a, n.ahead = 10)
plot(bv.vardec)
bv.vardec

# Residual diagnostics
#serial.test function takes the VAR model as the input.
serial.test(var.a)



#selecting the variables
# Granger test for causality
#for causality function to give reliable results we need all the
variables of the multivariate time series to be stationary.
causality(var.a, #VAR model
          cause = c("NEW")) #cause variable. If not specified the
n first column of x is used. Multiple variables can be used.

causality(var.a, #VAR model
          cause = c("AHE")) #cause variable. If not specified the
n first column of x is used. Multiple variables can be used.

causality(var.a, #VAR model
          cause = c("AGE")) #cause variable. If not specified the
n first column of x is used. Multiple variables can be used.

causality(var.a, #VAR model
          cause = c("REG")) #cause variable. If not specified the
n first column of x is used. Multiple variables can be used.

causality(var.a, #VAR model
          cause = c("UMP")) #cause variable. If not specified the
n first column of x is used. Multiple variables can be used.

causality(var.a, #VAR model
          cause = c("EMP")) #cause variable. If not specified the
n first column of x is used. Multiple variables can be used.

## Forecasting VAR models
```

```
fcast = predict(var.a, n.ahead = 12) # we forecast over a short h
orizon because beyond short horizon prediction becomes unreliable
or uniform
par(mar = c(2.5,2.5,2.5,2.5))
plot(fcast)


AntallMottakere = fcast$fcst[1]; AntallMottakere # type list


# Extracting the forecast column
x = AntallMottakere$NEW[,1]; x

tail(trainingdata)


# Inverting the differencing
#To get the data to the original scale we invert the time series
#since the values are just difference from the previous value, to
get the values on the original scale we add the last value from t
he DAX time series to the predicted values.
#the plot of the predicted values will also show that over longer
horizon the predicted values are not reliable
x = cumsum(x) + 1632
par(mar = c(2.5,2.5,1,2.5)) #bottom, left, top, and right
plot.ts(x)



#Linear Regression
mottakere <- ts(dplr$NEW, start = c(2011,1,1), end = c(2020,3,1),
frequency = 12)
ts.plot(mottakere)
training <- window(mottakere, end=c(2019,3))

a1 <-tslm(training~trend+ season)
linear<-forecast(a1, h=12)
plot(linear)
linear
linear1 <-c(1723,1453,1441,1310,1347,1685,1708,1783,1463,1647,184
8,1710)
# Accuracy Calculation
actual <- new_receivers[new_receivers[["dato"]]>"2019-03-01",]
actual <- actual$AntallMottakere

#Accuracy of VAR model
accuracy(x,actual)

#Accuracy of linear model
accuracy(linear1, actual)


#Plotting the accuracy
```

```
Var <- as.data.frame(x)
linear2 <- as.data.frame(linear1)
actual1 <- as.data.frame(actual)

v1 <-cbind(Var, linear2, actual1)

ja<-ts(v1,
       frequency = 12,
       start = c(2019, 4))
t<-ts.plot(ja, col=c(1:3))
t+ legend("bottomright", legend = c("VAR", "lm", "Actual"),col=c(
1:3), lty = 1)



#Forecasting the next 21 months

stnry = diffM(dplr) #difference operation on a vector of time ser
ies. Default order of differencing is 1.

# Creating a VAR model with vars
var.a <- vars::VAR(stnry
                   ,
                   p = 3,
                   type= "trend",
                   season = 12

)

fcast = predict(var.a, n.ahead = 21) # we forecast over a short h
orizon because beyond short horizon prediction becomes unreliable
or uniform
par(mar = c(2.5,2.5,2.5,2.5))
plot(fcast)


AntallMottakere = fcast$fcst[1]; AntallMottakere # type list


# Extracting the forecast column
x = AntallMottakere$NEW[,1]; x

tail(mymts)


# Inverting the differencing
x = cumsum(x) + 1403
par(mar = c(2.5,2.5,1,2.5)) #bottom, left, top, and right
plot.ts(x)

# Adding data and forecast to one time series
par(mfcol=c(1,1), cex=0.6)
AntallMottakereinv =ts(c(mymts[,1], x),
                       start = c(2011,1),
```

```
                          frequency = 12)

plot(AntallMottakereinv)
```