WORKING PAPER

Nowcasting Norwegian household consumption with debit card transaction data

NORGES BANK RESEARCH

17 | 2020

KNUT ARE AASTVEIT, TUVA MARIE FASTBØ, ELEONORA GRANZIERA, KENNETH SÆTERHAGEN PAULSEN AND KJERSTI NÆSS TORSTENSEN

NORGES BANK

Working papers fra Norges Bank, fra 1992/1 til 2009/2 kan bestilles over e-post:
FacilityServices@norges-bank.no

Fra 1999 og senere er publikasjonene tilgjengelige på www.norges-bank.no

Working papers inneholder forskningsarbeider og utredninger som vanligvis ikke har fått sin endelige form. Hensikten er blant annet at forfatteren kan motta kommentarer fra kolleger og andre interesserte. Synspunkter og konklusjoner i arbeidene står for forfatternes regning.

Working papers from Norges Bank, from 1992/1 to 2009/2 can be ordered by e-mail
FacilityServices@norges-bank.no

Working papers from 1999 onwards are available on www.norges-bank.no

Norges Bank's working papers present research projects and reports (not usually in their final form) and are intended inter alia to enable the author to benefit from the comments of colleagues and other interested parties. Views and conclusions expressed in working papers are the responsibility of the authors alone.

ISSN 1502-8190 (online)
ISBN 978-82-8379-174-7 (online)

Nowcasting Norwegian Household Consumption with Debit Card Transaction Data *

Knut Are Aastveit† Tuva Marie Fastbø ‡ Eleonora Granziera§

Kenneth Sæterhagen Paulsen¶ Kjersti Næss Torstensen ‖

November 8, 2020

Abstract

We use a novel data set covering all domestic debit card transactions in physical terminals by Norwegian households, to nowcast quarterly Norwegian household consumption. These card payments data are free of sampling errors and are available weekly without delays, providing a valuable early indicator of household spending. To account for mixed-frequency data, we estimate various mixed-data sampling (MIDAS) regressions using predictors sampled at monthly and weekly frequency. We evaluate both point and density forecasting performance over the sample 2011Q4-2020Q1. Our results show that MIDAS regressions with debit card transactions data improve both point and density forecast accuracy over competitive standard benchmark models that use alternative high-frequency predictors. Finally, we illustrate the benefits of using the card payments data by obtaining a timely and relatively accurate nowcast of the first quarter of 2020, a quarter characterized by heightened uncertainty due to the COVID-19 pandemic.

Keywords: Debit Card Transaction Data, Nowcasting, Forecast Evaluation, COVID-19

JEL classification: C22, C52, C53, E27

*This working paper should not be reported as representing the views of Norges Bank. The views expressed are those of the authors and do not necessarily reflect those of Norges Bank. We thank Todd Clark, John Galbraith, an anonymous referee from the Norges Bank Working Paper Series, seminar and conference participants at the 26th IIF Workshop, the 2020 Conference on Real-Time Data Analysis, Methods and Applications and the Bank of Canada.

†Norges Bank & BI Norwegian Business School; email: Knut-Are.Aastveit@Norges-Bank.no

‡Norges Bank; email: Tuva-Marie.Fastbo@Norges-Bank.no

§Norges Bank; email: Eleonora.Granziera@Norges-Bank.no

¶Norges Bank; email: Kenneth.Paulsen@Norges-Bank.no

‖Norges Bank; email: Kjersti-Ness.Torstensen@Norges-Bank.no

1 Introduction

Accurate knowledge of current economic conditions is essential for forecasting and economic policy making. However, key macroeconomic time series aggregates are released with a significant lag. Therefore, policy makers have to rely on early estimates of macroeconomic aggregates, or "nowcasts". These early assessments are often based on indicators which are highly correlated with the variable of interest, available at a higher frequency and released without delays (Evans, 2005; Giannone et al., 2008; Banbura et al., 2011).

The recent shutdown of significant portions of the worldwide economy, in order to restrain the outbreak of the coronavirus, has triggered a global recession. The uncertain consequences of the rapid spread of the virus and the induced infection control measures have made it extremely challenging for forecasters and policymakers to quantify and assess the current and future outlook of the economy. This has raised a renewed interest in the search for reliable high-frequency indicators that can track the real economy in a timely matter. Recent work by Carriero et al. (2020) and Lewis et al. (2020) have emphasized the importance of incorporating information from various weekly indicators, covering labour market conditions, the production side of the economy and financial markets, to track the real economy during the COVID-19 pandemic.[1] While these are clearly important indicators for tracking the real economy, there is a lack of reliable and timely indicators that capture the demand side of the economy, which may be of particular importance during the current pandemic. As highlighted in a recent paper by Guerrieri et al. (2020), economic shocks associated with the COVID-19 epidemic can be thought of as supply shocks that trigger changes in aggregate demand that are larger than the shocks themselves. They argue that when shocks are concentrated in certain sectors, as during a shutdown in response to an epidemic, there is greater scope for total spending to contract.[2]

In this paper, we document that debit card transaction data serve as an early and reliable indicator for household consumption in Norway. We use data for debit card transactions via BankAxept, provided by Nets Branch Norway (2006-2018) and Vipps AS (2018-2020), which records all domestic debit card transactions in physical terminals by Norwegian bank account holders.[3] The data is a proxy for household spending, is available without delays and without sampling errors. Debit cards are the dominant means of card payment for Norwegians, and it currently account for 9 out of 10 card transactions, and for more than 50% of the total value of all household consumption expenditures.[4] While this paper focuses on the case of Norway, we expect card transactions data to be useful for real time monitoring of consumption also in other countries, particularly for those where card payments account for a high share of consumption expenditures.

[1]The importance of using predictors that are sampled at a higher frequency than monthly observations, have also been earlier emphasized by Andreou et al. (2013) and Aastveit et al. (2017), using daily financial data and weekly activity and financial conditions indices, respectively.

[2]The standard mechanism is that when workers lose their income due to the shock they reduce their spending. However, what makes the shutdown situation so peculiar, is the fact that now some goods are no longer available making it less attractive to spend in general. Guerrieri et al. (2020) argue that the shutdown increases the shadow price of goods in affected sectors, making current consumption more expensive and thus discouraging it.

[3]This captures all domestic debit card transactions by Norwegian bank account holders, with the exception of online shopping.

[4]In the early years of our sample the shares was even higher, and debit cards transactions accounted for 55% of total consumption.

The card payment data are available at the weekly frequency, while consumption is sampled at the quarterly frequency.[5] To account for the frequency mismatch, we estimate Mixed Data Sampling (MIDAS) regressions which exploit the difference in sampling frequencies explicitly.[6] MIDAS models were introduced by Ghysels et al. (2004) as a flexible and parsimonious approach to relate low frequency data with high frequency variables.[7] The model specification is based on distributed lag polynomials and generates a direct forecast of the low-frequency variable. MIDAS models differ on the functional form of the polynomial characterizing the coefficients associated with the high frequency variables. Foroni et al. (2015) shows that if the difference in sampling frequency is not large, as in our case, it is not necessary to use distributed lag functions and that an unrestricted model performs comparable to the non-linear models, but has the advantage that it can be estimated with OLS. For these reasons, we prefer the Almon lag polynomial which restrict the number of parameters to estimate but does not require non-linear estimation.

In our empirical application, we compute recursive point and density nowcasts from the various MIDAS regressions for quarterly household consumption growth in Norway for the evaluation period 2011Q4-2020Q1. For the BankAxept data, we consider specifications that use both the volume of transactions as well as the value of payments as high-frequency predictors. We compare forecasts from the MIDAS models using BankAxept data with various alternative benchmark models, such as a simple AR model, a dynamic factor model and various MIDAS models that use alternative high-frequency predictors for consumption such as the unemployment rate, retail sales, car sales volumes, the purchase manager index (PMI), a business tendency survey for production of consumer goods, an uncertainty index[8] based on Norwegian newspaper data (Larsen, 2017), a Norwegian financial news index (FNI) (Larsen and Thorsrud, 2019; Thorsrud, 2020) and total returns from the benchmark index at the Oslo Stock Exchange. For monthly frequency predictors, we evaluate the forecasts performance of the models at three points in time during the quarter, just after the 1st, 2nd and 3rd month, respectively, while for weekly predictors we obtain a nowcast for each week in the quarter. In our evaluation, we focus on both point and density nowcasts. The relative forecasting performance of the different models is assessed in terms of root mean squared errors (RMSE) for point forecasts and the logarithmic score (LS) for the density forecasts.

We report three main findings. First, debit card transactions data are useful for nowcasting household consumption. Compared with the alternative benchmark models, we find gains both

[5]The data is available at the weekly frequency since January 2006 and at the daily frequency since January 2019. Due to the short sample of data available at the daily frequency we limit our analysis to the use of card payment data at the weekly frequency.

[6]Alternative approaches that also exploit differences in sampling frequencies range from bridge equation models (Angelini et al., 2011), factor models (Giannone et al., 2008; Banbura and Modugno, 2014), mixed-frequency VAR models (Kuzin et al., 2011; Schorfheide and Song, 2015) and structured machine learning regressions for high-dimensional time series data (Babii et al., 2020).

[7]MIDAS models are a popular way of dealing with mixed-frequency data and have been extensively used for nowcasting macroeconomic variables, see among others, Andreou et al. (2010, 2013), Clements and Galvão (2008, 2009), Kuzin et al. (2011), Marcellino and Schumacher (2010) and Ferrara and Marsilli (2019).

[8]Motivated by the large literature documenting that large changes in uncertainty may cause severe negative impacts on the economy, see e.g. Bloom (2009) and Baker et al. (2016), we include a Norwegian specific news-driven measure of uncertainty. In a recent paper, Baker et al. (2020) have argued that forward-looking uncertainty measures, such as a newspaper based uncertainty measure, are particularly useful for assessing the impact of the COVID-19 crisis on the economy.

in terms of improved point and density nowcasts performance, from MIDAS models that include debit card transactions data. The gains are sizeable and statistically significant. This holds true at all three points in time during the quarter for the monthly debit card data and starting from week 5 for the weekly data. In fact, while most of the alternative models are either performing at par or slightly worse than a simple AR model, models that incorporate BankAxept data provide gains in the magnitude of almost 60% compared with the simple AR model at the end of the quarter (approximately 6 weeks prior to the release of the quarterly national accounts).[9]

Second, for MIDAS models which contain BankAxept data, we find a gradual improvement of the nowcasting performance throughout the quarter, both for the volume and value predictors. Already after four weeks we obtain considerable gains in terms of improved nowcasts compared to the alternative benchmark models. In contrast, for the MIDAS models including other predictors the performance with respect to the AR model deteriorates in the second month, once the quarterly figures for the previous quarter consumption are released.

Finally, we document that a MIDAS model that includes the number of BankAxept transactions provide a strikingly accurate nowcast of the first quarter of 2020. As many other countries, Norway implemented drastic restrictions as a response to the coronavirus outbreak in the second week of March, two weeks after its first registered coronavirus case, including closing many shops and establishments. The shutdown had severe consequences for the economy where the unemployment rate rose from 2.7 percent in February to 10.7 percent at the end of March. Similarly, output and spending fell drastically. Despite such an extreme event, we show that a MIDAS model that includes the number of BankAxept transactions, provides a density nowcast at the very end of the first quarter that is almost centered around the actual value of -10.2 percent! As a comparison, the actual value of consumption for 2020Q1 either fall outside or lies in the tail of the density nowcasts provided from all the alternative benchmark models.

Our paper contributes to the large and growing literature on nowcasting. While most of the earlier papers focus on point nowcasts of GDP growth, we differ on two important aspects. First, instead of nowcasting GDP growth, we focus on household consumption growth. Household consumption consists of about 45 percent of GDP and has played a key role in the early parts of the current COVID-19 crisis. Second, we focus on both point and density nowcasts. Policymakers and forecasters are increasingly interested in forecast metrics that require density forecasts of macroeconomic variables, as complete probability distributions of outcomes provide information helpful for making economic decisions, see e.g. Tay and Wallis (2000), Garratt et al. (2003), Gneiting (2011) and Clark (2011). Accordingly, several central banks, including the Bank of England, Norges Bank and Sveriges Riksbank have committed to publishing density or interval forecasts for macroeconomic aggregates in recent years. However, despite the flourishing theoretical and empirical literature on the use of various mixed-frequency approaches for nowcasting, the focus has so far mainly been on point forecasts, Aastveit et al. (2014), Mazzi et al. (2014), Carriero et al. (2015), Aastveit et al. (2017) and Aastveit et al. (2018) being notable exceptions. The COVID-19 pandemic has triggered a massive spike in uncertainty, making the

[9]One exception is MIDAS models which includes monthly retail sales data. For these models, we obtain gains of about 50%, compared with the simple AR model, when data for all three months of the quarter are available. However, in contrast to the BankAxept data which have no publication delay, retail sales data have a publication delay of close to 1 month and are subject to revisions.

value of probabilistic forecasts more important than ever. As highlighted above, we document how our density nowcasts of household consumption provide important insights during the first quarter of 2020.

While most earlier studies in the nowcasting literature analyze the US economy, we focus on Norway. Nowcasting of the Norwegian economy has so far targeted output: Aastveit et al. (2011) describe Norges Bank's system for averaging models (SAM) which generates density forecasts for Norwegian Mainland GDP by combining vector autoregressive models, leading indicator models and factor models. Aastveit and Trovik (2012) estimate a dynamic factor model on a large database and find that unemployment, industrial production, and asset prices improve accuracy for the nowcasts of Norwegian GDP. Luciani and Ricci (2014) document that a Bayesian dynamic factor model outperforms simple univariate benchmark models both in terms of point and density forecast. Differently from this literature, our paper focuses on private household consumption which represents an important input to policy decision, given that it constitutes one of the main components of gross domestic product. Moreover, we explore the role of high frequency indicators for nowcasting.

The increased availability of electronic payments data has spurred a recent literature on real time forecasting of economic activity and its components using these electronic transactions. Duarte et al. (2017) obtain nowcast and one step ahead forecasts of Portuguese private consumption by combining data from ATM and POS terminals. Carlsen and Storgaard (2010) investigates whether electronic payments by card (Dankort) provides a useful indicator for the nowcast of monthly retail sales in Denmark. Verbaan et al. (2017) analyse whether the use of debit card payments data improves the accuracy of the nowcast and one quarter ahead forecast of Dutch private household consumption. Barnett et al. (2016) estimate a mixed frequency dynamic factor model which includes data on credit card transaction volumes to obtain a measure of US monthly GDP. Galbraith and Tkacz (2018) generate nowcast of Canadian GDP and retail sales using electronic payments data, including both debit card transaction and cheques clearing through the banking system. Finally, Aprigliano et al. (2019) assess the ability of a wide range of retail payment data to accurately forecast Italian GDP and its main domestic components. All these papers document an improvement in *point* forecast accuracy when using transaction data relative to simple benchmark models. Our paper differs from these studies in two important ways: first, we evaluate the performance of transaction data for both point and *density* forecasts and second we illustrate the use of transaction data during the highly uncertain and tumultuous environment generated by the COVID-19 pandemic. Moreover, differently from the electronic transaction data used in other studies, our dataset covers the universe of debit card transactions used by domestic card owners in domestic shops. Most other studies limit their analysis to transactions obtained from a single bank or from the settlement system for retail payments, which excludes 'on-us' payments, i.e. payments in which payer and payee hold accounts in the same bank.

Finally, we are related to the growing number of papers on forecasting economic aggregates during the COVID-19 pandemic. Carriero et al. (2020) focus on US GDP, Cascaldi-Garcia et al. (2020) on Euro-Area GDP. Lahiri and Yang (2020) nowcast U.S. state revenues. Aaronson et al. (2020) target US unemployment insurance, while Foroni et al. (2020) forecast GDP growth for

G7 countries. Hacioglu et al. (2020) use a large cross section of transaction data to nowcast consumption in the UK during the first half of 2020. Bounie et al. (2020) and Pascal et al. (2020) study households expenditure behavior during the pandemic with French electronic payment data and Dutch transaction data respectively.

The paper is organized as follows: section 2 describes the debit card transaction data, the other variables included in the analysis and the models used to conduct our investigation. Section 3 outlines the forecast evaluation results over the whole sample while section 4 illustrates the nowcast exercise during the COVID-19 pandemic. Section 5 concludes.

2 Data and Methodology

2.1 Debit Card Transaction Data

Norway is a near cashless economy: as shown in figure 1, the share of cash withdrawals relative to total card usage has fallen from 20 percent in 2011 to 8 percent in 2019.[10] This makes Norway an ideal economy in which to use electronic payments data to study household consumption.

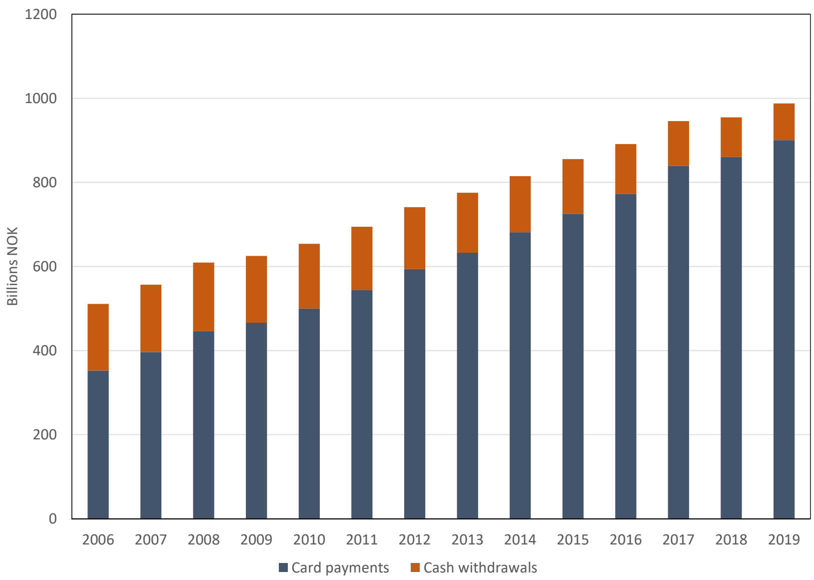


Figure 1: Value of card usage over time in Billions NOK. The blue and red bars show total value of card usage and cash withdrawals over time, respectively.

The debit card data are provided by the Norwegian retail clearing institution, Nets Branch Norway, for the years 2006 to 2018, and by Vipps AS for the years 2019 to 2020. The observations spanning the years 2006 to 2018 are sampled at the weekly frequency, whereas the observations for 2019 and 2020 are daily.[11] The data covers all debit card transactions via BankAxept, which is the national payment system in Norway owned by the Norwegian banks. Typically, all debit card payments in domestic physical terminals are BankAxept, whereas payments abroad, online

[10] Due to large fees charged by banks when taking out cash directly from an account, most cash withdrawals are done with cards.

[11] For the years 2006 to 2018 we can observe weekly payments by individuals classified by 26 categories and postal code of the shop, but we will only make use of aggregate transactions.

or mobile payments are paid through VISA or Mastercard. Note that in Norway online shopping represents only 3 percent of quarterly consumption. Our dataset has the advantage of including 'on-us' transactions, i.e. transactions completed by buyer and seller holding accounts in the same bank. These are usually omitted from data sets where transactions are recorded as they pass through the interbank payments system.

BankAxept debit card ownership and usage is well spread among age groups: figure 2 shows for different age groups the percentage of people that have completed at least one transaction using a BankAxept debit card in 2018. This percentage is close to 100 for all age groups. Therefore, our data is not capturing spending by a specific age group.

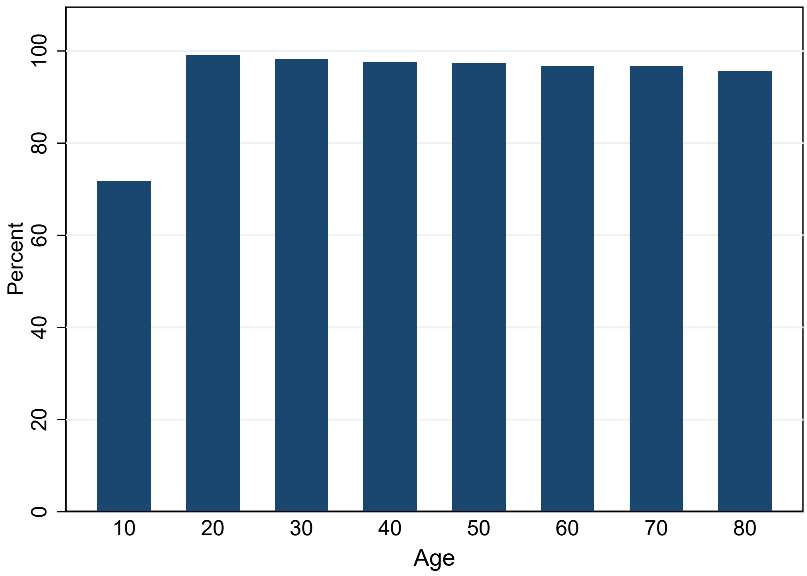
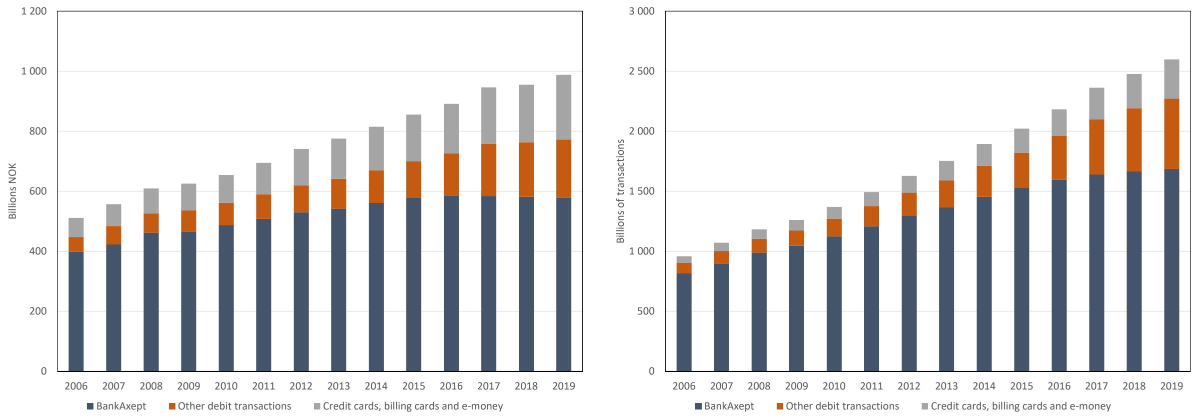


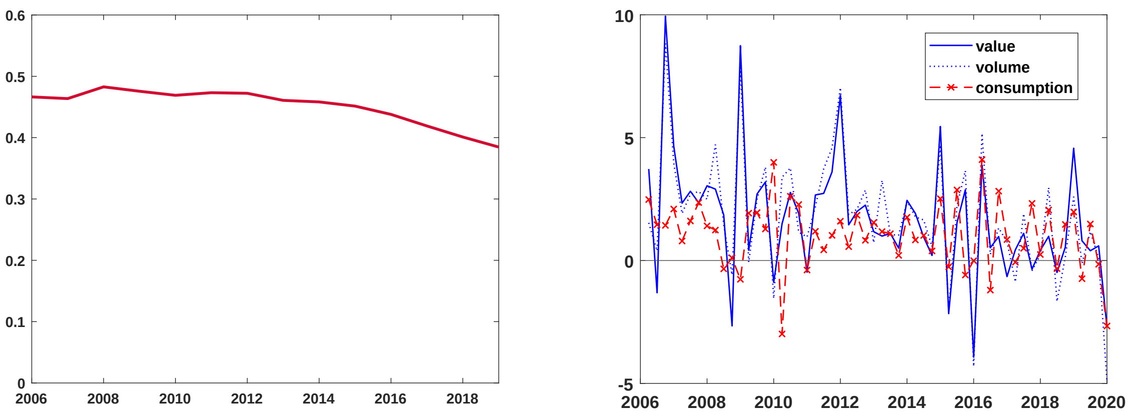
Figure 2: Percentage of population by age group that has completed at least one transaction using a BankAxept card in 2018. The horizontal axis reports the first year included in the age bracket.

Among card payments, debit card is the dominant means of payment in Norway throughout the sample period and account for 9 out of ten card transactions, see figure 3. On average from 2006 to 2019 eight out of ten card transactions were BankAxept, accounting for 71 percent of the total value. However, the share of BankAxept transactions has somewhat fallen over the sample period. In 2006 BankAxept accounted for 85 percent of the transactions and 78 percent of the value, whereas it accounted for 65 percent of the transactions and 58 percent of the value in 2019. Debit card data are used for smaller value transactions than credit card data, with an average value of 340NOK for BankAxept versus 640NOK for credit card. Unfortunately, credit card data are not available to us at high frequency for a substantial period. However, we do have information on the value of all card transactions, split between BankAxept, VISA and MasterCard, for the costumers of a large Norwegian bank from 2018 onwards. In that sample we observe that credit card and BankAxept data comove: the correlation between the month-over-month growth rate of value of BankAxpet and credit card transactions is 0.62. More importantly, the correlation between BankAxept value and total value of card transactions is 0.92, reflecting the fact that BankAxept is the dominant form of payment. We also find that credit card and BankAxept exhibit a similar behaviour during the first quarter of 2020, with the value of credit card data falling slightly more than debit cards: this is because while households completed more purchases online, they decreased their purchases abroad substantially. All these considerations lead us to conclude that the BankAxept data track total card usage very well.



(a) Value of Transactions (b) Volume of Transactions

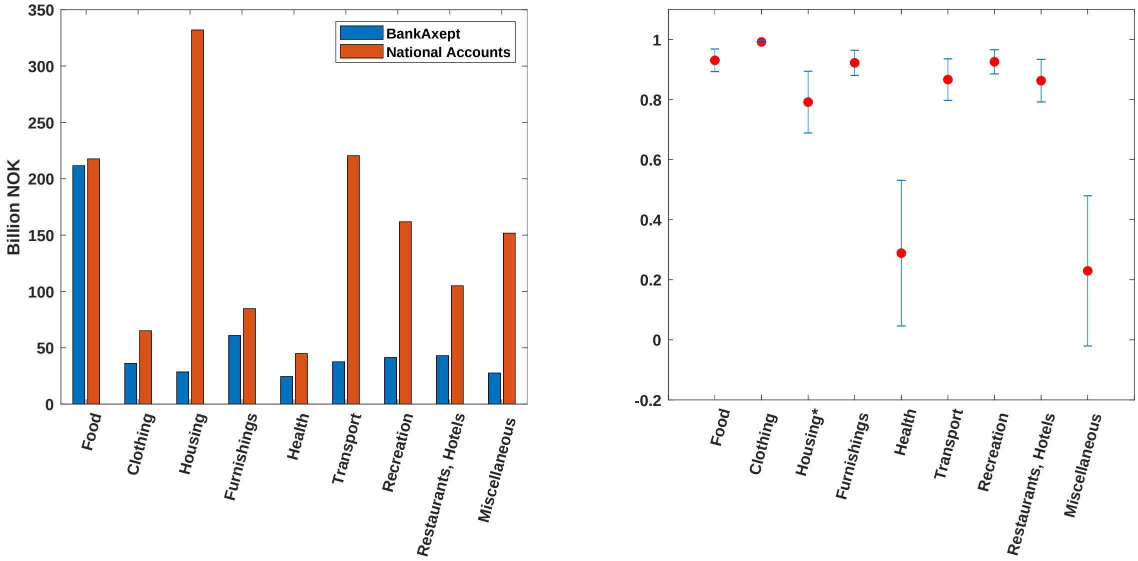
Figure 3: Card usage across categories over time. Panel A and B show the value of transactions, measured in Billions NOK, and the volume of transactions, measured in Billions of transactions, respectively, for three categories over time. The blue bars shows transactions made by BankAxept, the red bars shows other debit transactions and the grey bars shows transactions made by credit cards, billing cards and e-money.



(a) Share (b) Quarterly Growth Rates

Figure 4: (a) Debit card transactions as a share of total household consumption over time. (b) Quarter over quarter percentage changes of consumption, debit card values and number of debit card transactions over time

Still, the total value of debit card transactions (via BankAxept) represents about forty percent of total household consumption, as shown in figure 4, panel (a). Moreover, panel (b) of figure 4, which plots the quarter over quarter percentage changes of consumption, debit card values and number of debit card transactions, suggests that debit card transaction data track changes in consumption rather well. Therefore, at a first visual inspection, it seems promising to use these electronic payments data to nowcast consumer behaviour.



(a) Coverage

(b) Correlation

Figure 5: (a) Value of BankAxept Data and Consumption from National Accounts for several subcategories (Billion NOK): 'Food' refers to Food, beverages and tobacco, 'Clothing' to Clothing and footwear, 'Housing' to Housing, water and heating, 'Furnishing' to Furnishings and household equipment, 'Transport' to Cars and transport, 'Recreation' to Recreation and culture, 'Restaurants, Hotels' to Restaurants and hotels and 'Miscellaneous' to Miscellaneous goods and services. Values refer to the year 2019. (b) Correlation coefficient (absolute value) between Value of BankAxept Data and Consumption for several subcategories for quarterly data between 2006Q1 and 2020Q1, absolute value. Note that only the coefficient for the 'Housing' subcategory is negative.

The coverage of debit card data for different consumption components is quite heterogeneous. Figure 5 shows for the main consumption components the value of debit card transactions compared to the values of personal consumption from the National Accounts. BankAxept data cover well food and beverages, clothing and footwear and furnishing, but they are less successful in capturing housing, water and heating, transport, recreation and culture and restaurants and hotels. The latter is due to the fact that the BankAxept data does not include transactions by foreigners, which are included in the national accounts for subgroups, and that it is more common to pay hotels using credit cards. Overall, our data covers goods consumption well, while consumption of services is captured less precisely. Interestingly, despite the low coverage for housing, transport, recreation and restaurants and hotels, for these subcategories the correlation coefficient between the quarterly growth rate of the value of BankAxept data and the quarterly growth rate of the corresponding consumption subcategory is above 80%. Only for health and miscellaneous goods, which together represent about 4% of total household consumption, the correlation coefficient is less than 30%. Household consumption is made of 49% of services consumption and 47% of good consumption. The remaining 4%, which we do not capture with our debit card data, is given by the difference between direct purchase abroad by Norwegian resident households and direct purchase by non-residents in Norway.

The data is available at the daily frequency only from January 2019. Therefore, given the limited time span of observations at the daily frequency, we focus on data aggregated to the weekly frequency, i.e. the sum of the value of all transactions within a week.

Our main target variable is the quarterly Norwegian total household consumption, at current prices. The quarter over quarter growth rate of Norwegian household consumption has been relatively stable since 1990, averaging at 1.2% per quarter. As shown in panel (b) of figure 4, during the financial crisis consumption growth dropped for three consecutive quarters in 2009, following a decrease in economic activity at the time. The first quarter of 2020 has seen a severe drop in consumption, with the quarter over quarter growth rate plummeting to -2.7% for the seasonally adjusted series and to an astonishing -10.2% for the unadjusted series. For both series the decline is more severe than the drop registered during the financial crisis.

2.2 Alternative predictors

In addition to the debit card data, we consider several alternative predictors for consumption, capturing sales, financial conditions, uncertainty and labor market conditions. Specifically we include a retail sales index, car sales volumes, the purchase manager index, the benchmark index at the Oslo stock exchange, a financial news index based on textual data for Norway (Larsen and Thorsrud, 2019; Thorsrud, 2020), a macro uncertainty index based on textual data for Norway (Larsen, 2017), and the unemployment rate. These variables are available at a high frequency, early in the quarter of reference and, except for the retail sales index, are not subject to revisions.

We also consider the Business Tendency Survey for manufacturing, production of consumer goods, which measures the expected change in next quarter. This survey is available at the quarterly frequency and with about four weeks delay with respect to the reference quarter.

Name	**Transf**	**Timing**	**Publishing lag**
BankAxept	Δln	Every Monday	1 day
Stock Prices	Δln	Daily	1 day
Unemployment	*Level*	1st wd of month	1-3 days
Uncertainty	*Level*	Every Thursday	4 days
Car Sales	Δln	3-6th of month	3 to 6 days
Financial News	*Level*	1-3rd of month	1 month
PMI	*Level*	3-6th of month	1 month
Retail Sales	Δln	25-30th of month	1 month
BT Survey	*Level*	Every Quarter	1 month

Table 1: Variable names are reported in Column 1. Column 2 reports the transformation that is used for each of the explanatory variables in various MIDAS regressions. Finally, Column 3 and Column 4 indicate the official dates of the publication and the lag with which the data are reported, respectively.

2.3 Models Specification

In the first part of our analysis, we compare the forecasting accuracy of different models and predictors to nowcast consumption over the sample 2011Q4-2020Q1. This section describes the details of our baseline evaluation exercise.

As a benchmark model we consider a simple autoregressive (AR) model, which has proven to be a competitive benchmark:

$$y_t = \alpha + \sum_{p=1}^{P} \beta_p y_{t-p} + \varepsilon_t \tag{1}$$

where y_t is the quarter on quarter growth rate of consumption at the quarterly frequency, P the number of lags and ε_t are normally distributed errors.

Second, we consider models that include exogenous regressors. First, for quarterly available variables, we estimate:

$$y_t = \alpha + \sum_{k=0}^{K} \beta_k x_{n,t-k} + \varepsilon_t \tag{2}$$

where $x_{n,t}$ is the exogenous predictor n available at the quarterly frequency (like our business tendency survey) and ε_t are normally distributed errors.[12]

With the exception of the survey data, the predictors considered in our study are available at a higher frequency than the target variable y_t. Traditionally, the issue of mixed frequency data has been addressed by converting the higher-frequency data to the sampling rate of the lower frequency data, for example by temporally aggregating monthly indicators to quarterly, and adding it as regressors in (2). Clearly, this approach does not make use of the high frequency information available during the quarter of interest. To explicitly account for high frequency regressors, we use a mixed-data sampling (MIDAS) approach. The MIDAS model controls for parameter proliferation and is particularly suitable for our analysis given the size of our evaluation sample.

The general MIDAS model can be written as:

$$y_t = \alpha + \beta B\left(L^{1/m}; \theta\right) x_{n,t}^{(m)} + \varepsilon_t \tag{3}$$

where $x_{n,t}^{(m)}$ is variable n observed at the high frequency (e.g. $m = 3$ for monthly data and $m = 13$ for weekly data when the dependent variable is quarterly), $B\left(L^{1/m}; \theta\right) = \sum_{k=0}^{K} B\left(k; \theta\right) L^{k/m}$, $L^{1/m}$ is a lag operator such that $L^{k/m} x_{n,t}^{(m)} = x_{n,t-k/m}^{(m)}$ and the lag coefficient in $B\left(k; \theta\right)$ associated to the lag operator $L^{k/m}$ is parameterized as a function of a low-dimensional vector of parameters θ. As for the previous models we assume ε_t are normally distributed errors. MIDAS models differ according to the aggregation weighting scheme $B\left(k; \theta\right)$. In this paper we use the Almon lag polynomial, which has the advantage to retain linearity of the model in (3) with respect to the parameter vector θ. This aggregation scheme assumes that the lag weights are a linear function of M estimable nuisance parameters: $B\left(k; \theta\right) = \sum_{i=0}^{M} \theta_i k^i$. Following the literature, in our baseline specification we set $M = 2$ and we use the past three months of data to obtain the nowcast. Also, the predictors are the quarter over quarter log-difference for the variables in growth rates.

This MIDAS model is very flexible and is easily modified to include weekly rather than monthly debit predictors. In this case, we just need to set $m = 13$, which correspond to number of weeks in a quarter, while we can keep the size of the parameter vector θ fixed to $M = 2$.

[12] For both the model in (1) and in (2) the lag length is fixed to one: $P = 1$, $K = 1$. Including only one regressor and fixing the lag structure allows us to retain parsimony, which is important given the relatively small number of observations available, and to evaluate the performance of each single predictor.

Therefore, we are always using the past 13 weeks of data to nowcast. The regressors are the quarter over quarter log-difference of the variables in growth rates, e.g. the growth from week one in the first quarter of 2013 to week one in the second quarter of 2013.

The variable $x_{n,t}$ will either be one of the predictors described in Table 1, or a single factor computed from a larger pool of time series variables which are routinely used by Norges Bank staff to produce nowcasts for real economic activity (see Aastveit et al. (2011)). These include sub-indexes of industrial production, measures of employment and housing market indicators, all available at the monthly frequency.[13]

3 Real time nowcasting with debit card data

In this section, we evaluate the out-of-sample nowcasting performance of the debit card transaction data over the whole evaluation sample. In the next section, we focus on the first quarter of 2020 and show how debit card data provided an early signal of the sharp decline in consumption.

3.1 The design of the empirical analysis

In Norway, quarterly national accounts statistics are released by Statistics Norway six to seven weeks after the end of the reference quarter. We are interested in an estimate of the current quarter or the past quarter before the first release from Statistics Norway. For the models using monthly predictors, we evaluate the forecast performance of the models for three points in time, corresponding to the release of month 1, month 2 and month 3 information for the various predictors. For most predictors the month 1 data are available during the first week of month 2, prior to the release of national accounts for the pervious quarter, while month 2 and 3 data are available during the first week of month 3 and the first week after the end of the quarter, respectively. In the last two cases the national accounts for the pervious quarter has been released. For example, when nowcasting 2011Q4 we produce a first nowcast during the first week of November 2011, when the consumption figures for 2011Q3 are not yet available and we only observe the first month of the monthly indicators, a second nowcast during the first week of December 2011 after the release of national accounts for 2011Q3 and two months of monthly indicators, and finally a last nowcast during the first week of January 2012, when we have the full quarter information for the monthly indicators but not yet the release for 2011Q4. Note, however, that we give an informational advantage to the retail sale index, as this variable is normally released towards the end of the month with a one month lag. In our evaluation, we "pretend" that the retail sale index is released at the same time as the other predictors, i.e., during the first week of the month.

For the MIDAS model using weekly data we obtain a nowcast for each week in the quarter, starting from the second week, when data for the first week has just been released. The last nowcast is produced when the data for the thirteenth week has been released. Note that weekly data include transactions from Monday through Sunday.

[13] We have also extracted a factor from a large number of financial variables to compute a financial conditions index. However, given the relatively poor performance of this model, we do not report the results in our tables.

Nowcasts are obtained from models (1) through (3) in a recursive fashion with a rolling window estimation scheme.[14] Because of the relatively short size of the sample, we set the lag length of the autoregressive component to one. Table 1 describes the transformation of the predictors, as well as their highest available frequency and the release lag. The predictors are transformed as follows: BankAxept data, car sales, the retail sales index and stock prices are in quarter over quarter growth rates, while the uncertainty index, the financial news index, the PMI and the unemployment rate are in levels. As the debit card data are available starting in January 2006, the first estimation sample spans 2006Q1 to 2011Q3. Then, the evaluation sample runs from 2011Q4 to 2020Q1, for a total of 34 observations. The exercise is done using real time data for consumption and nowcasts are evaluated against the first data release. Following the literature, we evaluate point forecast accuracy by comparing root mean squared forecast errors (RMSFE), while for density forecasts we compute log-scores.

3.2 Out-of-sample nowcasting results

The results for point forecast accuracy are reported in Table 2, which shows the relative root mean squared error of the alternative models versus the AR model at the different forecast origins described above. A value smaller than one indicates that the alternative model performs better than the benchmark. The significance level for the equal predictive ability test by Diebold and Mariano (1995) are indicated by stars.[15] Note that the denominator is different as we move forward in time: up to week seven the AR nowcasts refer to the AR model where we do not know the value of the previous quarter consumption, because the national account statistics have not been released yet. For weeks eight to thirteen, the AR nowcasts make use of the knowledge of the past quarter consumption. Also, for data sampled weekly, i.e. BankAxept data, stock prices and uncertainty index, the nowcast is produced for each week in the quarter. For example, the row "week 1" reports the nowcast produced when the first week of data is available. For monthly sampled regressors, the nowcast is produced after the end of each month in the reference quarter, e.g. the row "week 5" reports the nowcast produced after the first month of data is released.

The table highlights that the AR is indeed a competitive benchmark, as for most models the relative RMSFE is larger than one and the difference in performance is statistically significant at all forecast origins considered. After the first month of data the relative RMSFE ranges from 1.046 for the car sales to 1.168 for the PMI predictor. The relative performance deteriorates over the quarter so that after the full quarter information is available, i.e. after three months for monthly data or after thirteen weeks for weekly data, the relative RMSFE ranges from 1.047 for car sales up to 1.476 for the uncertainty index. The monthly MIDAS models including the unemployment rate or the real activity factor improve over the benchmark, although only early in the quarter, when the previous quarter consumption figures are not available. Interestingly,

[14]We find the rolling estimation scheme preferable to the expanding window one given the increasing share of the value of debit card transactions over consumption. For robustness we report the results of the evaluation exercise using the expanding window in the appendix.

[15]The issue of real-time data complicates any assessment of whether the resulting differences in forecast accuracy between models are significant, see Clark and McCracken (2009). Monte Carlo evidence in Clark and McCracken (2015) indicates that, with nested models, the Diebold Mariano test compared against normal critical values can be viewed as a somewhat conservative (in the sense of tending to have size modestly below nominal size) test for equal accuracy in the finite sample.

the quarterly survey indicator has the smallest relative root mean squared error after the first month of the quarter, but the relative performance deteriorates drastically as we move along the quarter.[16] In contrast, gains for the BankAxept predictors, both volume and value, as well as for retail sales are sizable, statistically and economically significant and increasing over the quarter. After the first month of data the BankAxept value, unemployment rate and retail sales exhibit similar relative RMSFE, around 0.84 while the BankAxept volume improves even further with respect to the benchmark, showing gains up to 26%. After nine weeks or two months of data retail sales improves by 50% over the AR, while gains for the BankAxept models are still large, around 30%, but more modest. Finally, after the whole information for the quarter becomes available, i.e. at week thirteen for weekly data or after three months for monthly data, BankAxept value and retail sales provide gains of up to 55% while BankAxept volume of up to 45%.

Note that we are giving a notable advantage to the retail sales series for two reasons: first, although the series is revised, we are using as predictor the latest vintage available to date.[17] Second, and more importantly, we are assuming that the retail sales data are available without delays at the end of the reference period, while they are released with one month lag. This implies that the retail sales for the last month of the quarter are released just two weeks before the release of the national accounts data, while BankAxept data are released the day after the end of the reference month, i.e., 6-7 weeks prior to the release of the national accounts. These considerations are important for the assessment of the relative performance of the indicators: while comparable in terms of forecast accuracy, the BankAxept data are a much more timely indicator, as they are available about four weeks earlier, and they are not revised. Moreover, they are available at the weekly frequency. Overall, BankAxept data, both value and volume, are accurate predictors of consumption, providing large and statistically significant gains over a very competitive benchmark and alternative indicators. Also, gains improve over the quarter.

Table 3 shows the results for the density forecasts for the monthly and weekly frequency predictors. The same considerations as for point forecast accuracy hold for density forecasts. Consumption growth is a very good predictor of future consumption. For most MIDAS models the information included in the extra predictors does not help to improve over the AR model. Except for models including BankAxept data and retail sales, the relative performance of the models deteriorates over the quarter. Similar to the case of point nowcasts, the best models in terms of density nowcasting performance are the ones that include the Bank-Axept data, either value or volume, and the retail sales index for monthly indicators.

[16] Note that we are disregarding the fact that the survey data is available after the end of the reference quarter, so we are giving an unfair advantage to this indicator compared to the BankAxept data.

[17] Unfortunately, real time data for the retail sale index in Norway is not available.

Week	BA-Val	BA-Vol	Stock Prices	Uncert Index	Car Sales	Retail Sales	Financ News	Unemp. Rate	PMI	Real Act Factor	Survey
			Weekly				**Monthly**				**Quarterly**
1	0.951	1.137*	1.175**	1.132***							
2	1.066	0.998	1.146**	1.127***							
3	1.057	0.862**	1.139***	1.112***							
4	0.964	0.801**	1.147***	1.128***							
5	0.858**	0.738***	1.151***	1.165***	1.046	0.843***	1.153***	0.839***	1.168***	0.747***	0.693***
6	0.829***	0.806***	1.163***	1.156***							
7	0.788***	0.794***	1.192***	1.189***							
8	0.794**	0.799**	1.389***	1.338***							
9	0.700***	0.690***	1.394***	1.346***	1.020	0.494***	1.404***	1.021	1.400***	1.074**	0.800***
10	0.619***	0.618***	1.400***	1.412***							
11	0.507***	0.474***	1.405***	1.466***							
12	0.356***	0.368***	1.432***	1.439***							
13	0.456***	0.558***	1.471***	1.476***	1.047	0.438***	1.353***	1.186	1.385***	0.907	1.434***

Table 2: RMSFE relative to the AR model. The RMSFE for the AR model is 3.98 for week 1 through 7, when the previous quarter figures for consumption have not been released yet, and equals 3.47 for weeks 8 through 13 after the release of the previous quarter value. Entries smaller than one indicate the alternative model is more accurate than the AR. The differentials in the squared forecast errors are corrected for heteroskedasticity and autocorrelation using Newey and West (1987). '*', '*** and '***' indicate significance levels for the Diebold and Mariano (1995) test at the 10, 5 and 1 percent respectively. For weekly available data nowcasts are produced for each week in the quarter. For monthly predictors nowcasts are produced at three forecast origin: at week 5, when one month of information for the current quarter is available, at week 9 with two months of data for the reference quarter, at week 13 immediately after the end of the reference quarter, when all monthly information is available. Note that Retail Sales is available about one month after the end of the reference month.

Week	BA-Val	BA-Vol	Stock Prices	Uncert Index	Car Sales	Retail Sales	Financ News	Unemp. Rate	PMI	Real Act Factor	Survey
			Weekly				**Monthly**				**Quarterly**
1	-2.815	-2.939*	-2.977**	-2.942***							
2	-2.892	-2.791	-2.956***	-2.937***							
3	-2.863	-2.626***	-2.961***	-2.926***							
4	-2.756	-2.539***	-2.980***	-2.943***							
5	-2.626***	-2.445***	-2.990***	-2.973***	-2.930**	-2.605**	-2.959***	-2.684	-2.951	-2.625**	-2.513*
6	-2.596***	-2.569***	-3.007***	-2.963***							
7	-2.565**	-2.599*	-3.586***	-3.476***							
8	-2.410**	-2.418**	-3.591***	-3.362***							
9	-2.262***	-2.245***	-3.575***	-3.477***	-2.934	-1.932***	-3.789***	-2.761	-3.721***	-2.863*	-2.808
10	-2.215**	-2.235**	-3.569***	-3.588***							
11	-2.530	-2.228**	-3.584***	-3.740***							
12	-2.087*	-1.991***	-3.729***	-3.673***							
13	-1.966***	-2.089***	-3.933***	-3.779**	-3.031	-2.162***	-4.168***	-2.844*	-3.749***	-2.644	-3.372

Table 3: Log Scores, averages over the evaluation sample. '*', '*** and '***' indicate significance levels for the Amisano and Giacomini (2007) test at the 10, 5 and 1 percent respectively. The average log score for the AR model is -2.822 for week 1 through 7, when the previous quarter figures for consumption have not been released yet, and equals -2.688 for weeks 8 through 13 after the release of the previous quarter value.

3.3 Distinguishing between goods consumption and consumption of services

Figure 5 in Section 2.1 showed that the BankAxept data covered goods consumption well, while consumption of services was captured less precisely. To investigate further what are the source of the improvements for the BankAxept data over the benchmark, we split the BankAxept predictors in two subcategories, goods and services. We then repeat the forecast evaluation exercise for total household consumption as described above. Results are reported in table 4. The results reveal that the value and volume of goods are more accurate predictors of household consumption than the value and volume of services. In fact the value of services provide limited gains over the benchmark AR model, while the volume of services does not outperform the benchmark.

Week	Value Goods	Value Services	Volume Goods	Volume Services
1	0.891*	0.978	1.131*	1.103*
2	1.034	1.029	1.034	1.030
3	1.040	1.052*	0.879**	1.004
4	0.932	1.082***	0.776***	1.011
5	0.818***	1.101***	0.689***	1.038
6	0.803***	1.061	0.756***	1.097
7	0.772***	0.967	0.755***	1.134*
8	0.781***	0.983	0.771***	1.241**
9	0.681***	0.898*	0.659***	1.205*
10	0.605***	0.867**	0.597***	1.188*
11	0.531***	0.631***	0.505***	0.864
12	0.393***	0.859	0.343***	1.199
13	0.466***	0.955	0.487***	1.096

Table 4: RMSFE relative to the AR model for total household consumption. Target variable: qoq growth rate of total household consumption. RMSFE for the AR model: 3.98 for week 1 through 7 and 3.47 for weeks 8 through 13. Entries smaller than one indicate the alternative model is more accurate than the AR. The forecast errors are corrected for heteroskedasticity and autocorrelation using Newey and West (1987). '*', '*** and '***' indicate significance levels for the Diebold and Mariano (1995) test at the 10, 5 and 1 percent respectively

To explore further the distinction of subcomponents of the debit card data (goods vs services), we investigate the ability of the two subcomponents to nowcast goods consumption and services consumption in the national accounts. The relative RMSFE for this experiment are shown in Table 5. Unsurprisingly, the value and volume of BankAxept data for goods transactions are accurate predictors of goods consumptions, with gains over the benchmark model that reach 80% at week 12. The value of services outperforms the AR model but gains are limited and mostly not statistically significant. Interestingly, the value of goods is as accurate as the value of services for nowcasting services consumption. Moreover, starting from week 8 the volume of goods exhibits a lower RMSFE than the volume of services. Overall, the results from this exercise and from Table 4 suggest that while the value of debit card data for goods delivers an accurate nowcast for total consumption as well as for both goods and services consumption, the value of debit card data for services is a reliable indicator only of consumption of services.

	Goods Consumption				Services Consumption			
Week	Value Goods	Value Services	Volume Goods	Volume Services	Value Goods	Value Services	Volume Goods	Volume Services
1	0.576***	0.735***	0.976	1.075*	1.309***	1.436***	1.344*	1.289***
2	0.799***	0.816***	0.976	1.038	1.274**	1.392***	1.179	1.135*
3	0.916**	0.859**	0.905***	1.029	1.105	1.353***	1.035	1.098
4	0.896***	0.921*	0.842***	1.024	0.964	1.296***	0.962	1.094
5	0.832***	0.976	0.780***	1.032	0.908	1.207**	0.955	1.128
6	0.810***	0.977	0.827***	1.061**	0.898	1.112	0.937	1.200
7	0.793***	0.918***	0.816***	1.122***	0.862	1.022	0.943	1.239
8	0.777**	0.891***	0.784***	1.170***	0.635***	0.694***	0.695***	0.862
9	0.696***	0.806***	0.675***	1.129**	0.678***	0.733***	0.736***	0.915
10	0.585***	0.759***	0.571***	1.115**	0.745***	0.781**	0.785**	0.927
11	0.392***	0.566***	0.373***	0.874	0.844***	0.819*	0.876*	0.884
12	0.197***	0.846	0.210***	1.158	0.716***	0.696***	0.754***	0.735***
13	0.366***	0.959	0.516***	1.122**	0.671***	0.652***	0.721***	0.803*

Table 5: RMSFE relative to the AR model for goods consumption (left 4 columns) and services consumption (right 4 columns). Target variable: qoq growth rate of goods or services consumption. RMSFE for the goods consumption AR model: 8.904 for week 1 through 7 and 8.294 for weeks 8 through 13. RMSFE for the services consumption AR model: 1.606 for week 1 through 7 and 2.309 for weeks 8 through 13. Entries smaller than one indicate the alternative model is more accurate than the AR. The forecast errors are corrected for heteroskedasticity and autocorrelation using Newey West (1987). '*', '*** and '***' indicate significance levels for the DM test at the 10, 5 and 1 percent respectively

3.4 Retail Sales vs BankAxept

We have already stressed the advantages of BankAxept data over the retail sales index in a real-time nowcasting setting, i.e. they are sampled at a higher frequency and without a publication lag. However, once the full quarter information is available for both variables, MIDAS models with BanAxept data and MIDAS models with the retail sales index exhibit fairly similar nowcasting performance. This could mean that the BankAxept data and the retail sales index carry much of the same information. While retail sales data are of course easy to obtain, the debit card data are more costly to collect and use for a practitioner, but they have huge timeliness advantages. We therefore investigate the marginal value of the debit card data above and beyond the retail sales data.

First, we focus on the in-sample analysis. In the left panel of Table 6 we show the estimated coefficients from a single regressor model where consumption is regressed on quarterly aggregated BankAxept Value, BankAxept volume and retail sales. We compare those coefficients with the ones from models where we augment the BankAxept value and BankAxept volume with retail sales. In all cases the estimated coefficients are positive and significant. In the augmented models, in which both retail sales and BankAxept data are included, the coefficients for retail sales are reduced with almost 30 percent while the coefficients for BankAxept are more than halved, reflecting some overlapping informational content. However, they remain highly statistically significant. The in-sample analysis therefore suggests that the two indicators are not only carrying the same information.

In-Sample Coefficients		**Out of Sample Relative RMSFE**			
Model	Full Quarter	Model	One Month	Two Months	Full Quarter
Single Regressor				**Relative to Benchmark**	
Ba-Val	0.343***	BA-Val, RS	0.546***	0.463***	0.406***
BA-Vol	0.506***	BA-Vol, RS	0.588***	0.468***	0.419***
RS	0.496***				
Multiple Regressors				**Relative to Retail Sales**	
BA-Val	0.112***	BA-Val, RS	0.648***	0.937	0.927*
RS	0.360***	BA-Vol, RS	0.698***	0.948	0.955
				Relative to BankAxept	
BA-Vol	0.140***	BA-Val, RS	0.542***	0.713***	0.874**
RS	0.393***	BA-Vol, RS	0.573***	0.692***	0.761***

Table 6: Left Panel: regression coefficients for quarterly variables over the full sample. '*', '*** and '***' indicate significance levels for the DM test at the 10, 5 and 1 percent respectively. Right Panel: relative RMSFE from MIDAS models with monthly predictors. "BA-Val, RS" ("BA-Vol, RS") refers to a MIDAS model with both BAX-Value (BAX-Volume) and Retail Sales as regressors. Behnchmark models are (i) AR; (ii) MIDAs monthly model with retail sales (iii) MIDAS monthly model with BankAxept;

Second, we turn our focus to the relative out-of-sample forecasting performance of the single predictor regression models versus a model that includes data on both BankAxept and retail sales. To ease comparison between models, for this exercise we consider MIDAS models where all predictors are sampled at the same frequency, i.e., at the monthly frequency. The right panel of table 6 shows the results. Adding retail sales to the models that only include BankAxept (either value or volume) as a predictor reduces RMSFE substantially and significantly at all forecast origins, although gains decline gradually over the quarter. Similar considerations hold if we add BankAxept data to a model including only retail sales, although in this case the marginal improvement is more limited and not statistically significant at the end of the quarter.

Overall, this analysis shows that even in an unfeasible setting in which retail sales and BankAxept were released on the same date, BankAxept data would provide additional gains over the retail sales index.

3.5 Robustness

We perform several robustness checks to corroborate the validity of our results: first, with respect to the use of final vintage data; second, with respect to results being driven by outliers; third with respect to the data transformation; finally with respect to the estimation scheme.

The choice of benchmark vintage is a key issue in any application using real-time vintage data (see Croushore (2006) for a survey of forecasting with real-time macroeconomic data). To make sure that our results are not driven by the use of first release data, we also evaluate the nowcast against the final data vintage. Results are reported in Table 7 in the Appendix. This experiment confirms the findings of our baseline analysis: the AR model is a hard benchmark to beat. From week 5 the BankAxept data, both value and volume are significantly more accurate than the benchmark. Towards the end of the quarter, the gains with respect to the benchmark exceed 40%.

Given the relatively short size of the evaluation sample, one might wonder whether the superior performance of the BankAxept data is driven by a few observations. To ease this concern, we plot the time series of the difference in the cumulative Root Mean Squared Forecast Errors between the AR model and each BankAxept model in Figure 10 in the Appendix for three forecast origins: after week 4, week 9 and week 12. The figure shows that the models including BankAxept data perform better than the AR throughout the evaluation sample, indicating that our results are not driven by a few extreme observations.

Our target variable is the seasonally unadjusted quarter over quarter growth of household consumption expenditure. The BankAxept data and other regressors are also seasonally unadjusted. We prefer using seasonally unadjusted data given the challenges associated with seasonally adjusting weekly data with a relatively short sample of data. One might be concerned that results might be different if we consider seasonality patterns. Therefore, as a further robustness check we report our results for the year over year growth rate of consumption in Table 8 of the Appendix. Regressors are also transformed into year over year growth rates. Although the gains of the BankAxept models with respect to the AR model are somewhat less impressive (up to 23% for the value and 35% for volume at week 13), they are the most accurate predictors of consumption from week 5 onwards. In fact retail sales, usually the most competitive alternative indicator, improves only by 8% once the full month information is available. Interestingly, stock prices outperform the benchmark, but only by 10% at week 13.

Finally, we check whether the results are robust to the estimation scheme. In our baseline analysis we use the rolling scheme which keeps the number of observations in the estimation sample constant. Our choice is motivated by concerns regarding changes in the parameters, given that the share of consumption explained by BankAxept data increases over time. The window size is relatively short as it includes only 22 observation and might results in high parameter uncertainty. Therefore, we try using the expanding window estimation scheme rather than the rolling one, with the final estimation sample including 55 observations. As shown in Table 9, results are unchanged.

4 Real Time Nowcasting During the COVID-19 Pandemic

In this section we illustrate the performance of our models for nowcasting the first quarter of 2020, characterized by increased uncertainty due to the coronavirus pandemic.

Norway implemented drastic restrictions as a response to the coronavirus outbreak on March 12th, two weeks after its first registered coronavirus case, including closing kindergartens, schools, many shops and establishments. Initially, the restrictions were put in place for two weeks and subsequently extended for three additional weeks. The lock-down was gradually lifted from late April, with kindergartens, primary schools and most establishments allowed to reopen, but distancing and other infection prevention measures still negatively affecting profitability.

As a result of the corona crisis, registered unemployment rose sharply in March. During the month more than 238,000 new unemployment benefit applications were registered. Consequently, the seasonally adjusted unemployment rate increased substantially at the end of the month to 10.7 percent, up from 2.7 percent in February. Stock prices plummeted by almost 25 percent and the uncertainty index almost doubled from February to March. At the same time the value of debit card transactions fell by 14 percent with respect to March 2019 and by 10 percent with respect to the previous month, while the volume of transactions dropped by 25 percent with respect to March 2019 and by 21 percent with respect to February 2020. In contrast, retail sales have fallen by only 6 percent in year over year terms, and have even increased by 3.7 percent with respect to February 2020. Consumption dropped by 10.2 percent in quarter-over-quarter terms and by 6% in year-over-year terms. The decline was driven mainly by a fall in goods consumption, which plummeted by 16.2% in quarter over quarter terms, while services consumption declined by 2%.

Figure 6 shows the time series of the weekly debit card data during the first quarter of 2020 compared with the same weeks in the first quarter of 2019 and figure 7 reports the evolution of the data for two sub-categories: goods and services. The figures document several interesting facts. First the volume of transactions dropped one week earlier than the value of payments and by a larger proportion. In fact the volume of transactions declined the week before the lockdown was implemented. This suggests that households were making fewer purchases even before the lockdown but the amount per transaction was larger. Second, both the value and volume of services started to plummet the week before the lockdown, as households were limiting social interactions in restaurants, bars and other establishments, reached the minimum the week after the beginning of the lockdown and did not recover in the sample considered. Finally, the value of goods increased dramatically during week eleven, when the lockdown was implemented and household stocked food items, declined sharply on week twelve and finally started to recover on week thirteen with the increase in value more pronounced than the increase in volume.

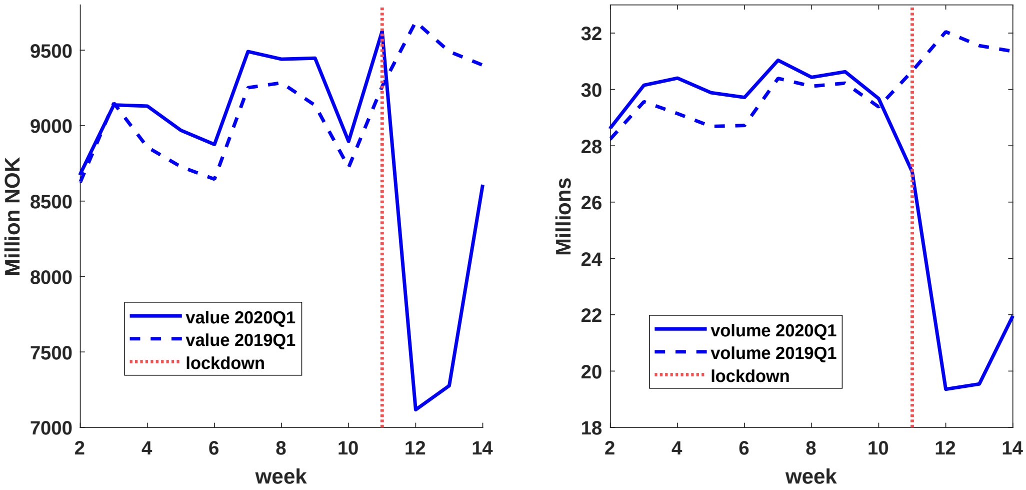


Figure 6: Weekly Debit Card Data 2020Q1: Value and Volume

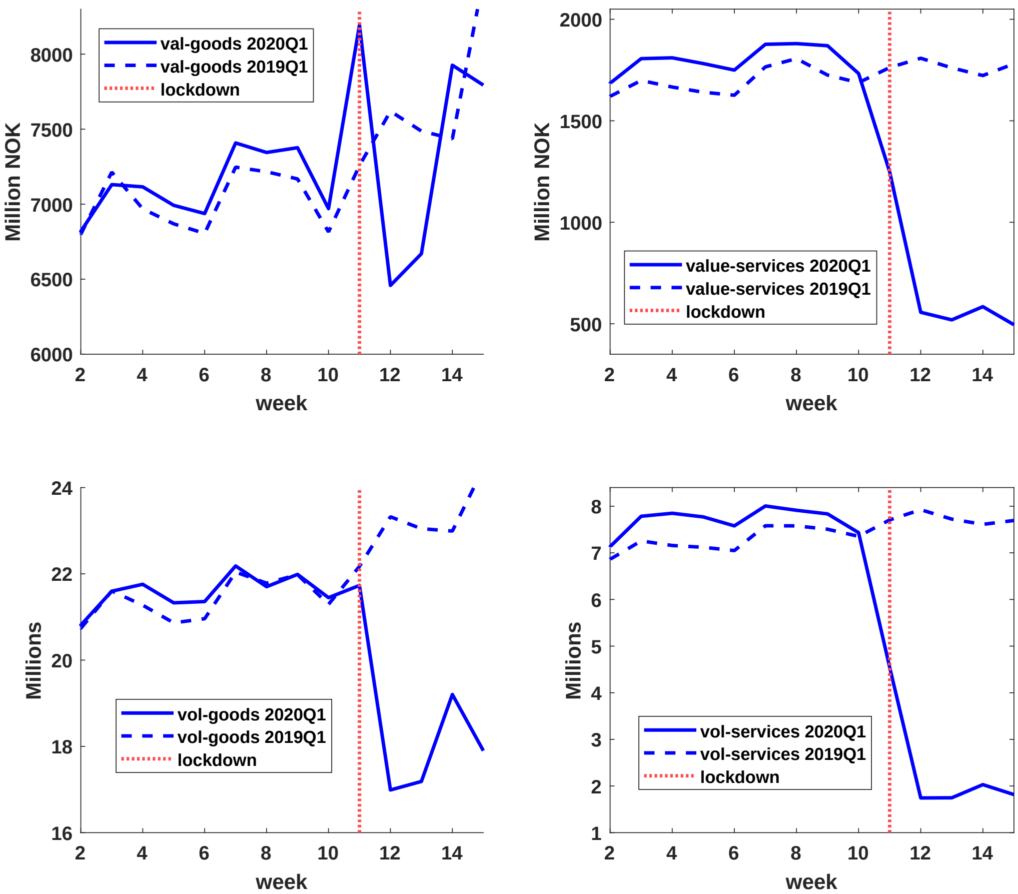


Figure 7: Weekly Debit Card Data 2020Q1 by Sub-Cathegories: Goods (left panels) and Services (right panels)

Although BankAxept is a very good indicator of household consumption in "normal" business cycles, this might not be the case during Covid-19. Due to the risk of getting infected, households partly shifted from physical stores to online shopping which our data does not cover. On the other hand, households reduced both domestic and international travels, due to domestic and foreign restrictions, and hence consumption of services typically paid with credit cards fell significantly. Using information from a large Norwegian bank on card payments via both debit and credit cards, we find that credit cards fell slightly more compared to BankAxept in 2020Q1 (QoQ-growth), but the fall in total card payments is about the same as the fall in debit card payments via BankAxept.

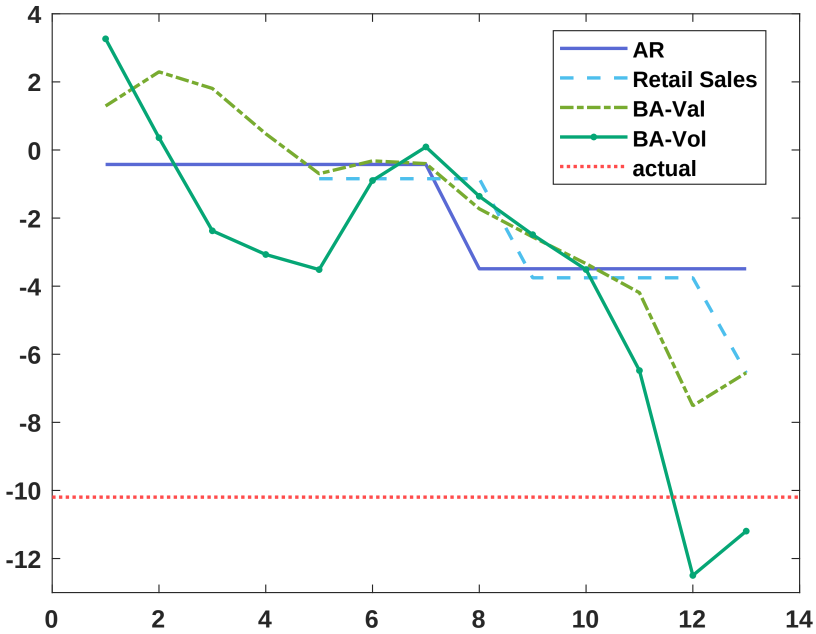


Figure 8: Nowcasting 2020Q1 household consumption: weekly point forecasts

We now evaluate the weekly performance of the BankAxept indicators during the first quarter of 2020 and the first week of the second quarter of 2020. The point forecast for 2020Q1 obtained for the AR model, the MIDAS models using retail sales and the BankAxept data are shown in Figure 8.[18] The AR model predicts a slightly negative growth rate of consumption early on in the quarter and the nowcast drops to -3.6% after the quarterly figure of consumption for the previous quarter has been released. The nowcast from the MIDAS model with retail sales is available from week 5 and initially almost coincides with the nowcast from the AR model. From week 9 to 12 the retail sales nowcast drops to the value obtained from the AR model. Finally, at the end of the quarter, after all three months of retail sales data are released, the nowcast falls further to -6%. The nowcast obtained from the weekly MIDAS models using BankAxept data is initially positive but declines gradually throughout the quarter. In particular from week

[18] In the figure we do not report the point forecasts from the MIDAS models including the other predictors, given the poor performance compared to the benchmark in the evaluation exercise of the previous section. The nowcast for 2020Q1 are quite inaccurate and range from 2% to -2% at the end of the quarter, after all monthly information becomes available.

11 onwards the nowcast falls below the ones from the retail sales and AR model and the highest drop is seen in week twelve for both value and volume. The prediction from the BankAxept volume model is lower than for the BankAxept value, reflecting a larger drop in the volume than in the value of transactions. On week 13 the nowcast form the MIDAS model with BankAxept volume is -10.7%, which is remarkably close to the actual value of -10.2%.

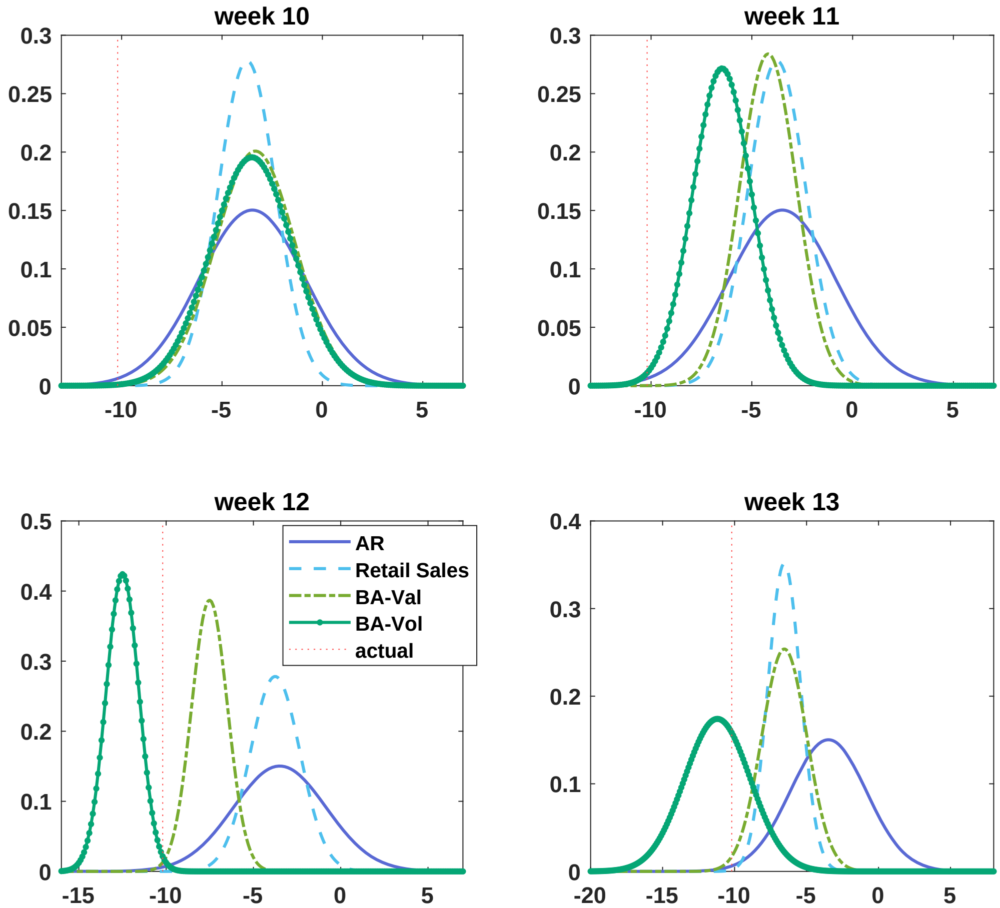


Figure 9: Nowcasting 2020Q1 household consumption: weekly density forecasts for the BanAxept MIDAS model and the AR model starting on week 10, which is the week preceding the lockdown.

Next, we show the density forecast for the MIDAS models with BankAxept data, the MIDAS model with retail sales data and the AR model for selected weeks in Figure 9. In week ten, the week before the lockdown measures were implemented, the models have a remarkably similar point forecast but the density from the AR model is more dispersed and therefore assigns a larger probability of a -10.5 percent fall in consumption. The density for the retail sales model is quite concentrated around the point forecast and assigns a negligible probability of observing a fall larger than 9%. The densities for the BankAxpet models, both value and volume are almost identical and their variance lies in between the one of the AR and the retail sales model. During the following week the densities for the weekly MIDAS models shift to the left, and the change is larger for the model with the volume of transactions. The densities become more concentrated around the mean. In week twelve we observe a further shift to the left for both BankAxept models and a further reduction in the variance. Finally, on week thirteen both the densities shift slightly to the right and become less concentrated around the point forecast. While clearly, the BankAxept volume model provides the more accurate nowcast, the BankAxept value is the second best performing model and assigns a higher probability than the AR or retail sales model of observing the actual realization for 2020Q1 consumption growth.

Although the evaluation exercise on the overall sample seemed to highlight a similar forecast accuracy of the retail sales and BankAxept data, the illustration of the 2020Q1 nowcast has proven that BankAxept data might be a more reliable and timely indicator during periods of large adverse shocks. Overall, our analysis shows that BankAxept data are a valuable indicator of household spending, not only in the longer evaluation sample but also in the highly uncertain environment generated by the pandemic.

5 Conclusion

In this paper we showed that debit card data are an accurate and reliable predictor of household consumption not only on average but also during periods of high uncertainty brought by large exogenous and unanticipated shocks. This data is available without delays and sampling errors and at a higher frequency than other commonly used indicators such as retail sales.

We documented sizable gains for point and density forecast over commonly used benchmark models to nowcast consumption and report that the improvements are statistically significant starting early in the quarter, when about one month of debit card data is available.

While this paper focus on the case of Norway, we expect debit card transactions data to be useful for real time monitoring of consumption also in other countries, particularly for those where card payments account for a high share of consumption expenditures. Although the data for parts of the sample period is available at the individual level, we have only used aggregated series in our analysis. One possible way to exploit the full richness of our data for forecasting, is to apply the techniques by Babii et al. (2020) and Babii et al. (2020) on structured machine learning regressions for high-dimensional time series data. We leave this for future research.

References

Aaronson, D., B. S., B. A., S. D., and B. Seo (2020). Pandemics, Huricanes, and Google Trends, UI:Forecasting Unemployment Insurance During Covid-19. mimeo.

Aastveit, K. A., C. Foroni, and F. Ravazzolo (2017). Density Forecasts With Midas Models. *Journal of Applied Econometrics 32*(4), 783–801.

Aastveit, K. A., K. Gerdrup, and A. S. Jore (2011). Short-term Forecasting of GDP and Inflation in Real Time: Norges bank's System for Averaging Models. Staff Memo 2011/9, Norges Bank.

Aastveit, K. A., K. Gerdrup, A. S. Jore, and L. A. Thorsrud (2014). Nowcasting GDP in Real Time: A Density Combination Approach. *Journal of Business and Economic Statistics 32*(1), 48–68.

Aastveit, K. A., F. Ravazzolo, and H. K. van Dijk (2018). Combined density Nowcasting in an uncertain economic environment. *Journal of Business & Economic Statistics 36*(1), 131–145.

Aastveit, K. A. and T. Trovik (2012). Nowcasting Norwegian GDP: the Role of Asset Prices in a Small Open Economy. *Empirical Economics 42*(1), 95–119.

Amisano, G. and R. Giacomini (2007). Comparing Density Forecasts via Weighted Likelihood Ratio Tests. *Journal of Business and Economic Statistics 25*(2), 177–190.

Andreou, E., E. Ghysels, and A. Kourtellos (2010). Regression models with mixed sampling frequencies. *Journal of Econometrics 158*(2), 246–261.

Andreou, E., E. Ghysels, and A. Kourtellos (2013). Should Macroeconomic Forecasters Use Daily Financial Data and How? *Journal of Business & Economic Statistics 31*(2), 240–251.

Angelini, E., G. Camba-Mendez, D. Giannone, L. Reichlin, and G. Rünstler (2011). Short-term forecasts of euro area gdp growth. *Econometrics Journal 14*(1), C25–C44.

Aprigliano, V., G. Ardizzi, and L. Monteforte (2019). Using Payment System Data to Forecast Economic Activity. *International Journal of Central Banking 15*(4), 55–80.

Babii, A., R. T. Ball, E. Ghysels, and J. Striaukas (2020). Machine learning panel data regressions with an application to nowcasting price earnings ratios. arXiv preprint arXiv:2008.03600.

Babii, A., E. Ghysels, and J. Striaukas (2020). Machine learning time series regressions with an application to nowcasting. arXiv preprint arXiv:2005.14057.

Baker, S. R., N. Bloom, and S. J. Davis (2016). Measuring Economic Policy Uncertainty. *The Quarterly Journal of Economics 131*(4), 1593–1636.

Baker, S. R., N. Bloom, S. J. Davis, and S. J. Terry (2020). COVID-Induced Economic Uncertainty. NBER Working Papers 26983, National Bureau of Economic Research, Inc.

Banbura, M., D. Giannone, and L. Reichlin (2011). Nowcasting. In M. P. Clements and D. F. Hendry (Eds.), *The Oxford Handbook of Economic Forecasting*. Oxford Handbooks in Economics.

Banbura, M. and M. Modugno (2014). Maximum Likelihood Estimation Of Factor Models On Datasets With Arbitrary Pattern Of Missing Data. *Journal of Applied Econometrics 29*(1), 133–160.

Barnett, W., M. Chauvet, D. Leiva-Leon, and L. Su (2016). Nowcasting Nominal GDP with the Credit-Card Augmented Divisia Monetary Aggregates. MPRA wp n. 73246.

Bloom, N. (2009). The Impact of Uncertainty Shocks. *Econometrica 77*(3), 623–685.

Bounie, D., Y. Camara, and J. Galbraith (2020). Consumers' Mobility, Expenditure and Online-Offline Substitution Response to COVID-19: Evidence from French Transaction Data. Working Paper 2020S-28, CIRANO.

Carlsen, M. and P. E. Storgaard (2010). Dankort Payments as a Timely Indicator of Retail Sales in Denmark. Danmarks Nationalbank Working Papers 2010-66.

Carriero, A., T. E. Clark, and M. Marcellino (2015). Realtime nowcasting with a Bayesian mixed frequency model with stochastic volatility. *Journal of the Royal Statistical Society Series A 178*(4), 837–862.

Carriero, A., T. E. Clark, and M. Marcellino (2020). Nowcasting Tail Risks to Economic Activity with Many Indicators. Working Papers 202013, Federal Reserve Bank of Cleveland.

Cascaldi-Garcia, D., G. D., M. M., and T. R. T. Ferreira (2020). Back to the Present: Learning about the Euro Area through a Nowcasting Model. mimeo.

Clark, T. E. (2011). Real-time density forecasts from bayesian vector autoregressions with stochastic volatility. *Journal of Business & Economic Statistics 29*(3), 327–341.

Clark, T. E. and M. W. McCracken (2009). Tests of Equal Predictive Ability With Real-Time Data. *Journal of Business & Economic Statistics 27*(4), 441–454.

Clark, T. E. and M. W. McCracken (2015). Nested forecast model comparisons: A new approach to testing equal accuracy. *Journal of Econometrics 186*(1), 160–177.

Clements, M. P. and A. B. Galvão (2008). Macroeconomic forecasting with mixed-frequency data. *Journal of Business & Economic Statistics 26*, 546–554.

Clements, M. P. and A. B. Galvão (2009). Forecasting US output growth using leading indicators: an appraisal using MIDAS models. *Journal of Applied Econometrics 24*(7), 1187–1206.

Croushore, D. (2006). *Forecasting with Real-Time Macroeconomic Data*, Volume 1 of *Handbook of Economic Forecasting*, Chapter 17, pp. 961–982. Elsevier.

Diebold, F. X. and R. Mariano (1995). Comparing predictive accuracy. *Journal of Business Economics and Statistics 13*(3), 253–263.

Duarte, C., P. M. M. Rodrigues, and A. Rua (2017). A Mixed Frequency Approach to the Forecasting of Private Consumption with ATM/POS Data. *International Journal of Forecasting 33*, 61–75.

Evans, M. D. (2005). Where are we now? real-time estimates of the macro economy. *International Journal of Central Banking 1*(2), 127–175.

Ferrara, L. and C. Marsilli (2019). Nowcasting global economic growth: A factor-augmented mixed-frequency approach. *The World Economy 42*(3), 846–875.

Foroni, C., M. M., and D. Stevanovic (2020). Pandemics, Huricanes, and Google Trends, UI:Forecasting Unemployment Insurance During Covid-19. CIRANO Working Papers 2020s-32.

Foroni, C., M. Marcellino, and C. Schumacher (2015). U-MIDAS: MIDAS Regressions with Unrestricted Lag Polynomials. *Journal of the Royal Statistical Society 178*(1), 57–82.

Galbraith, J. and G. Tkacz (2018). Nowcasting with Payment System Data. *International Journal of Forecasting 34*, 366–376.

Garratt, A., K. Lee, M. H. Pesaran, and Y. Shin (2003). Forecast uncertainties in macroeconomic modeling: An application to the UK economy. *Journal of the American Statistical Association 98*(464), 829–38.

Ghysels, E., P. Santa Clara, and R. Valkanov (2004). The MIDAS Touch: MIxed DA Sampling Regression Model. mimeo, Chapel Hill.

Giannone, D., L. Reichlin, and D. Small (2008). Nowcasting: The real-time informational content of macroeconomic data. *Journal of Monetary Economics 55*(4), 665–676.

Gneiting, T. (2011). Making and evaluating point forecasts. *Journal of the American Statistical Association 106*(494), 746–762.

Guerrieri, V., G. Lorenzoni, L. Straub, and I. Werning (2020). Macroeconomic Implications of COVID-19: Can Negative Supply Shocks Cause Demand Shortages? NBER Working Papers 26918, National Bureau of Economic Research, Inc.

Hacioglu, S., D. Känzig, and P. Surico (2020). Consumption in the time of Covid-19: Evidence from UK transaction data. Discussion Paper 14646, CEPR.

Kuzin, V., M. Marcellino, and C. Schumacher (2011). MIDAS vs. mixed-frequency VAR: Nowcasting GDP in the euro area. *International Journal of Forecasting 27*(2), 529–542.

Lahiri, K. and C. Yang (2020). State Revenue Forecasting in the Aftermath of COVID-19 Shutdown using Machine Learning. mimeo.

Larsen, V. H. (2017). Components of uncertainty. Working Paper 2017/5, Norges Bank.

Larsen, V. H. and L. A. Thorsrud (2019). The value of news for economic developments. *Journal of Econometrics 210*(1), 203–218.

Lewis, D., K. Mertens, and J. H. Stock (2020). U.S. Economic Activity During the Early Weeks of the SARS-Cov-2 Outbreak. NBER Working Papers 26954, National Bureau of Economic Research, Inc.

Luciani, M. and L. Ricci (2014). Nowcasting Norway. *International Journal of Central Banking 10*(4), 215–248.

Marcellino, M. and C. Schumacher (2010). Factor MIDAS for Nowcasting and Forecasting with Ragged-Edge Data: A Model Comparison for German GDP. *Oxford Bulletin of Economics and Statistics 72*(4), 518–550.

Mazzi, G., J. Mitchell, and G. Montana (2014). Density nowcasts and model combination: nowcasting euro-area gdp growth over the 2008-9 recession. *Oxford Bulletin of Economics and Statistics 76(2)*, 233–256.

Newey, W. and K. West (1987). A simple, positive semidefinite, heteroskedasticity and autocorrelation consistent covariance matrix. *Econometrica 55(3)*, 703–708.

Pascal, G., G. Kapetanios, N. Neuteboom, F. Ritsema, and A. Ventouri (2020). Consumption during the covid-19 pandemic: Lockdown or fear? Evidence from transaction data for the Netherlands. Working Paper 2020/04, King's Business School.

Schorfheide, F. and D. Song (2015). Real-Time Forecasting With a Mixed-Frequency VAR. *Journal of Business & Economic Statistics 33*(3), 366–380.

Tay, A. and K. F. Wallis (2000). Density Forecasting: A Survey. *Journal of Forecasting 19*, 235–254.

Thorsrud, L. A. (2020). Words are the new numbers: A newsy coincident index of the business cycle. *Journal of Business & Economic Statistics 38*(2), 393–409.

Verbaan, R., W. Bolt, and C. van der Cruijsen (2017). Using Debit Card Payments Data for Nowcasting Dutch Household Consumption. DNB Working Paper 2017-571.

6 Appendix

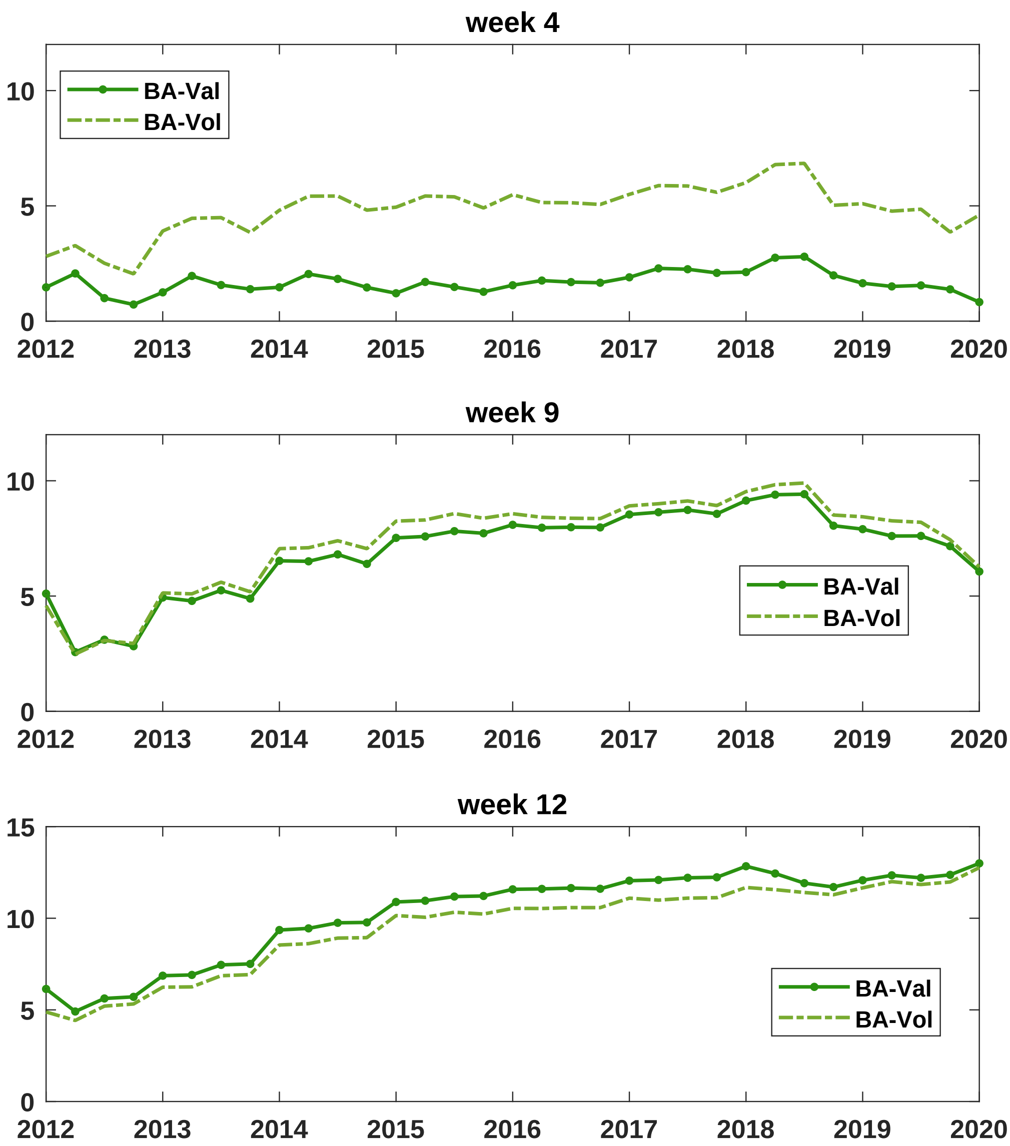


Figure 10: Time series of difference in cumulative Root Mean Squared Forecast Errors for selected weeks. At each point in time "t" the difference is defined as the cumulative RMSFE of the AR model minus the cumulative RMSFE for the BankAxept model up to time "t". Target variable is quarter over quarter growth of consumption.

Week	BA-Val	BA-Vol	Stock Prices	Uncert Index	Car Sales	Retail Sales	Financ News	Unemp. Rate	PMI	Real Act Factor	Survey
		Weekly						**Monthly**			**Quarterly**
1	0.902*	1.111*	1.114**	1.121***							
2	1.019	0.997	1.123***	1.098***							
3	1.028	0.876*	1.119***	1.099***							
4	0.962	0.834**	1.129***	1.107**							
5	0.877**	0.782***	1.131***	1.138***	0.984	0.812***	1.103***	0.780***	1.123***	0.743***	0.719***
6	0.857**	0.852**	1.143***	1.139***							
7	0.834**	0.854**	1.162***	1.176***							
8	0.857**	0.865*	1.344***	1.296***							
9	0.775***	0.753***	1.350***	1.323***	0.943	0.646***	1.337***	0.977	1.348***	1.061**	0.828***
10	0.688***	0.664***	1.356***	1.369***							
11	0.579***	0.528***	1.362***	1.405***							
12	0.558***	0.557***	1.388***	1.358***							
13	0.622***	0.685***	1.428***	1.395**	1.006	0.556***	1.288***	1.195*	1.358***	0.969	1.377***

Table 7: RMSFE relative to AR model, final release. The RMSFE for the AR model is 4.25 up to week 7 and 3.77 from week 8 onwards after the release of the previous quarter value. Entries smaller than one indicate the alternative model is more accurate than the AR. Differential in squared forecast errors are corrected for heteroskedasticity and autocorrelation using Newey and West (1987). '*', '*** and '***' indicate significance levels for the Diebold and Mariano (1995) test at the 10, 5 and 1 percent respectively. Predictors in growth rates are computed as qoq.

Week	BA-Val	BA-Vol	Stock Prices	Uncert Index	Car Sales	Retail Sales	Financ News	Unemp. Rate	PMI	Real Act Factor	Survey
		Weekly					**Monthly**				**Quarterly**
1	1.160***	1.285**	1.095*	1.020							
2	1.189	1.351	1.126*	0.996							
3	1.029	1.058	1.141*	0.984							
4	0.921	0.911	1.150*	1.000							
5	0.849	0.814	1.151**	0.980	1.181***	1.018	1.115***	1.063*	1.042*	1.082***	1.428***
6	1.115***	1.072	1.132**	1.011							
7	1.198*	1.095	1.080*	1.067							
8	0.937	0.865	1.031	1.018							
9	0.970	0.863	0.999	1.004	1.019	1.024	1.034	1.027	0.985	1.024	1.133***
10	0.860	0.816	0.955	1.097							
11	0.881	0.792	0.909	1.243*							
12	0.873	0.667	0.897	1.190*							
13	0.770	0.646	0.938	1.355*	1.172***	0.927	1.041	1.150*	0.934	1.010	1.210***

Table 8: RMSFE relative to the AR model for year-over-year growth rates of consumptions. Regressors are also computed as year-over-year growth. The RMSFE for the AR model is 1.36 for week 1 through 7, when the previous quarter figures for consumption have not been released yet, and equals 1.41 for weeks 8 through 13 after the release of the previous quarter value. Entries smaller than one indicate the alternative model is more accurate than the AR. The differentials in the squared forecast errors are corrected for heteroskedasticity and autocorrelation using Newey and West (1987). '*', '*** and '***' indicate significance levels for the Diebold and Mariano (1995) test at the 10, 5 and 1 percent respectively. For weekly available data nowcasts are produced for each week in the quarter. For monthly predictors nowcasts are produced at three forecast origin: at week 5, when one month of information for the current quarter is available, at week 9 with two months of data for the reference quarter, at week 13 immediately after the end of the reference quarter, when all monthly information is available. Note that Retail Sales is available about one month after the end of the reference month.

Week	BA-Val	BA-Vol	Stock Prices	Uncert Index	Car Sales	Retail Sales	Financ News	Unemp. Rate	PMI	Real Act Factor	Survey
		Weekly					**Monthly**				**Quarterly**
1	0.868***	1.019	1.066*	1.050***							
2	0.961	0.922*	1.065**	1.040*							
3	0.954	0.796***	1.061***	1.037*							
4	0.878***	0.738***	1.057***	1.038							
5	0.786***	0.671***	1.052***	1.062*	0.907***	0.766***	1.019	0.784***	1.069***	0.667***	0.637***
6	0.757***	0.720***	1.049***	1.043							
7	0.727***	0.708***	1.061***	1.059*							
8	0.762***	0.750***	1.255***	1.235***							
9	0.668***	0.665***	1.256***	1.246***	0.969	0.530***	1.246***	0.928	1.307***	1.029	0.754***
10	0.592***	0.602***	1.259***	1.276***							
11	0.502***	0.507***	1.263***	1.303***							
12	0.411***	0.456***	1.266***	1.281***							
13	0.488***	0.611***	1.271***	1.336**	0.921	0.470***	1.236***	1.135**	1.264***	0.888**	1.297***

Table 9: RMSFE relative to the AR model, all models estimated with the expanding window scheme. The RMSFE for the AR model is 4.30 for week 1 through 7, when the previous quarter figures for consumption have not been released yet, and equals 3.64 for weeks 8 through 13 after the release of the previous quarter value. Entries smaller than one indicate the alternative model is more accurate than the AR. The differentials in the squared forecast errors are corrected for heteroskedasticity and autocorrelation using Newey and West (1987). '*', '*** and '***' indicate significance levels for the Diebold and Mariano (1995) test at the 10, 5 and 1 percent respectively. For weekly available data nowcasts are produced for each week in the quarter. For monthly predictors nowcasts are produced at three forecast origin: at week 5, when one month of information for the current quarter is available, at week 9 with two months of data for the reference quarter, at week 13 immediately after the end of the reference quarter, when all monthly information is available. Note that Retail Sales is available about one month after the end of the reference month.