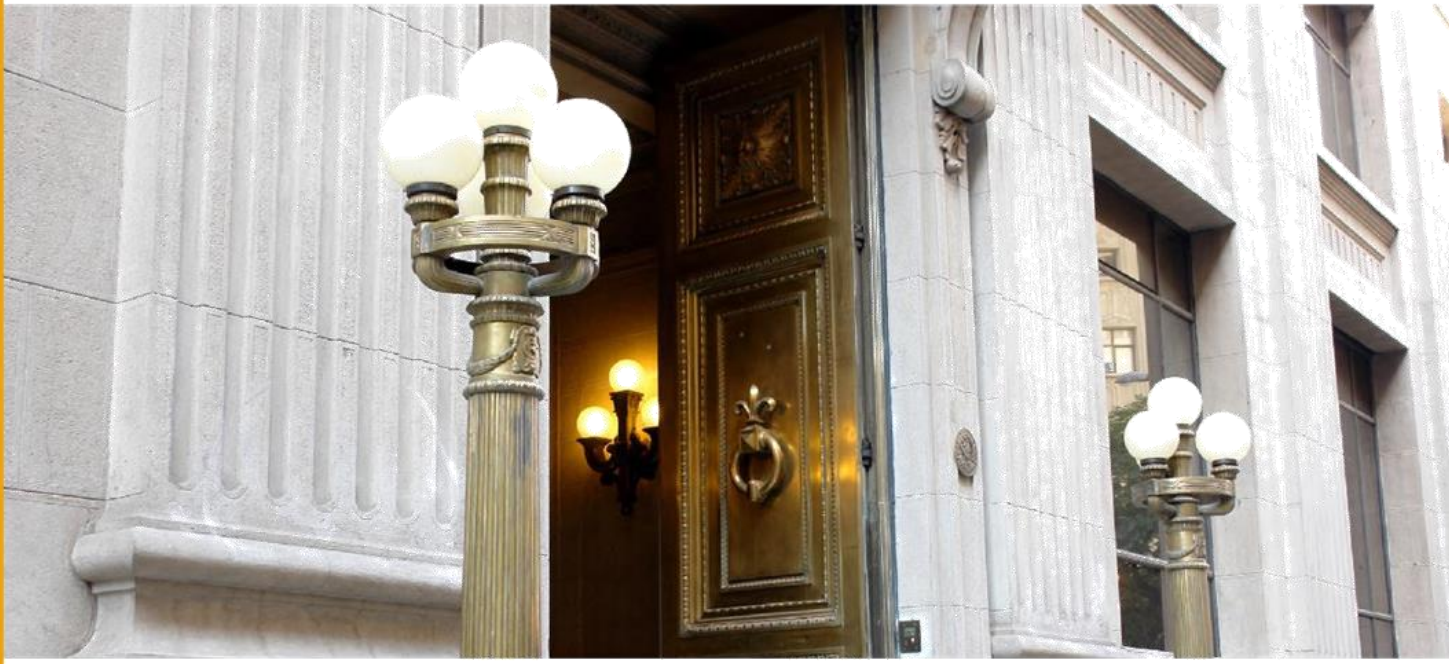
DOCUMENTOS DE TRABAJO

Nowcasting Chilean household consumption with electronic payment data

Marcus P. A. Cobb

N° 931 Noviembre 2021
BANCO CENTRAL DE CHILE



Documentos de Trabajo del Banco Central de Chile
Working Papers of the Central Bank of Chile
Agustinas 1180, Santiago, Chile
Teléfono: (56-2) 3882475; Fax: (56-2) 38822311

Documento de Trabajo
N° 931

Working Paper
N° 931

Nowcasting Chilean household consumption with electronic payment data

Marcus P. A. Cobb[§]
Central Bank of Chile

Abstract

When economies are hit by relevant shocks, the need to be able to follow developments in real-time increases for policymakers and private agents alike. When an event of this type is underway, the situation can change dramatically in a matter of days. The COVID pandemic is only the latest example. Electronic payment data is available with virtually no time lag and could therefore contribute to increasing the speed at which assessments are made. This paper makes use of a novel database to track Chilean household consumption in real time during the pandemic and compares the results to those of standard nowcasting methods used at the Central Bank of Chile. The results suggest that payment data is most useful as the shocks occur, when traditional models may have a harder time interpreting the information. The gain in more stable times is less obvious. The results also show, as one might expect, that the relationship between this naturally-occurring data and the variable of interest can be affected by its own shocks. In this case at least, electronic payments showed sudden shifts in intensity that needed to be accounted for in order to produce the final forecasts. All in all, the models based on payment data appear to be a relevant addition to the forecasting toolkit.

Resumen

Cuando la economía se ve afectada por perturbaciones relevantes, los hacedores de políticas y agentes privados necesitan estar informados sobre la situación con la mayor celeridad posible. Estos eventos inusuales pueden resultar en cambios relevantes de un día para otro. La pandemia causada por el COVID es sólo el último ejemplo. La información de medios de pagos electrónicos está disponible prácticamente sin rezago y, por lo tanto, podría contribuir a reducir el tiempo necesario para evaluar la situación. Este documento utiliza información de medios de pagos electrónicos para seguir la evolución del consumo de hogares en Chile en tiempo real durante el periodo de pandemia y compara los resultados con los de modelos tradicionales. El análisis sugiere que los datos de medio de pago son particularmente útiles en el momento que las perturbaciones ocurren cuando los modelos tradicionales presentan las mayores dificultades para interpretar la información. Su contribución en tiempos normales no es tan evidente. Los resultados también sugieren que las relaciones estadísticas entre los datos de medios de pago y las variables de interés pueden cambiar rápidamente por lo que es necesario estar atento a estas situaciones. Con todo, estos modelos parecen una adición relevante al conjunto de herramientas disponibles para hacer proyecciones.

§ Email: marcus.cobb@gmail.com. The views expressed here are solely those of the author and do not necessarily reflect the views of the Central Bank of Chile. I am grateful to an anonymous referee for some very useful comments. Any mistakes that remain are my responsibility.

1. Introduction

Policymakers and other institutions have a constant need for up-to-date information on the economic situation in order to make their decisions in an informed manner. Unfortunately, broad macroeconomic indicators, like GDP and its components, are usually not immediately available. These, in particular, are published with a considerable lag, are subject to revisions and are produced with a quarterly frequency at best. Because of this, the current assessment and forecasts for these variables need to be as accurate as possible. If forecasts deviate too much from realised values, real-time decisions may turn out to be wrong, causing inefficiencies and potentially avoidable cost. In this context, this paper explores the usefulness of electronic payment data in predicting Chilean household consumption in the current trimester. It focuses on the performance of a relatively simple nowcasting scheme over the last two years, which is when the COVID pandemic caused a major shock to the Chilean economy and the rest of the world.

Policy makers and forecasters strive to develop alternative methods that lower forecasting errors, so it is not surprising that, with the increased availability of high frequency data, the literature on concurrent period forecasting, or nowcasting, has evolved rapidly over the last few years. This strand focuses on using readily available data to predict the value in the current period for an indicator of interest that is not yet available for some reason. GDP is a typical example of one of these indicators of interest because of the considerable lag with which it is published. As regards to the readily available data, electronic payment data has received much attention lately, but others exist, for example energy demand, expectation surveys or internet search indicators. A quick search on the internet will provide an ample list of hits that will surely contain satisfying motivation and literature review for nowcasting. Because of the goal of this paper, Conesa et.al. (2015), Galbraith and Tkacz (2018) or Aastveit et.al. (2020) seem likely to be a good reference, due to their focus solely on payment data and their in-depth exposition of the characteristics of this type of information. Another is Gil et.al. (2019) who provide an overview in a context of other high-frequency data like indicators based on internet search queries. The most relevant result of the literature review presented in these references is that most of the applications find that payment data can contribute towards increasing accuracy when nowcasting broad economic indicators like GDP and household expenditure.

These papers, however, were written before the outbreak of COVID. Citing Aastveit et.al. (2020):

> "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."
>
> "…, 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."

As mentioned in Aastveit et.al. (2020) the COVID pandemic is a global phenomenon, meaning that probably everywhere where payment data is available, this has been used in an effort to improve nowcasting. They list a number of papers that basically do the same thing in different countries, all of which find it to be very useful. It should not come as a surprise that one of the objectives of this paper is to assess whether this is also true for Chile. The answer to this is yes. In other words, I find that electronic payment data serve as an early and reliable indicator for household consumption. However, Chile has certain idiosyncrasies that make the circumstances surrounding the pandemic somewhat different from those of other countries in the references. First, cash payments are still a relatively important share of expenditure, so having electronic payment data may not guarantee *a priori* reliable estimates. Second, over the last three years COVID has not been the only shock that has affected the Chilean economy. First there was an episode of social unrest in the last quarter of 2019. Then in March 2020 COVID hit. Then, in July of the same year, as a way of providing quick income to families in need, a "one time only" withdrawal from the individual mandatory contribution pension fund was approved by congress. Then a second one was approved in December and a third in May 2021. Currently a fourth one is under discussion. All in all, the three withdrawals amount to US$ 50 billion (equivalent to roughly 20% of GDP of 2020!!!). This string of unprecedented shocks provides a unique setting to put the robustness of the payment data models to the test.

In this context, I explore the nowcasting power of electronic payment data for Chile during the COVID pandemic by constructing an out-of-sample exercise in which I compare bridge equation models based solely on successive accumulation of such data against dynamically selected SARIMA models and some nowcasting models that use an array of monthly indicators. The period under study covers the last 3 years, from the last quarter of 2018 to the second quarter of 2021. Because of the limited number of observations of the evaluation period, the analysis is conducted as an events study as opposed to a general assessment. The main finding is that the bridge models based on payment data are more informative in periods that are subject to sudden shocks. The rest of the time they are not necessarily better than the broader models. Both these facts make the case for using the payment data models in conjunction with other models.

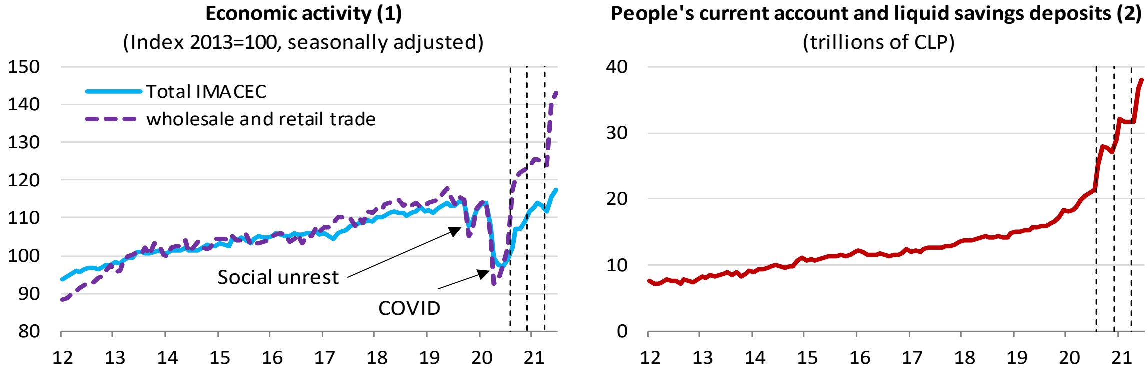
The paper is organized as follows: section 2 describes the expenditure and electronic payment data and the models used to evaluate the usefulness of the latter. Section 3 outlines the nowcasting exercise. Section 4 presents the overall conclusion and other remarks.

2. Data and Methodology

This section describes the characteristics of the data and the reasoning for arriving at the particular nowcasting framework. The context in which the models were developed was a couple of months after COVID hit, when many models were struggling to deal with the unpredictable movements in most of the indicators they used. To put the shocks from the last three years into perspective, figure 1 presents in its left panel the evolution of the Monthly Indicator of Economic Activity (IMACEC) over the last 10 years. After a relatively stable growth (2.4% average yearly growth Jan.13 to Sep.19) an outburst of social unrest in October 2019 meant that activity contracted around 3.5% y/y in that month and then due to COVID the following year in June 2020 it contracted approximately 14% y/y. The subsequent recovery meant that a year later, in June 2021, economic activity has exceeded pre-COVID levels (3.5% w/r to June 2019, meaning a close to 2% average yearly growth). As regards wholesale and retail trade, the numbers are more impressive. Social unrest produced a contraction of about 8% y/y, while COVID meant a maximum contraction of about 21% y/y and a year later the activity in this sector had increased a whopping 23% over the level of June two years earlier (about 11% average yearly growth).

As becomes clear from the figures, the situation is unique as of the third trimester of 2019. It is likely that commerce will continue to show strong numbers given that families are holding large sums of money in very liquid assets due primarily to the three early withdrawals from the individual mandatory contribution pension funds that were mentioned before. The magnitude of these holdings can be seen in the right panel of figure 1. All this provides the background to look at how household consumption has evolved recently and how well electronic payment data helps to predict it.

Figure 1: Evolution of key variables



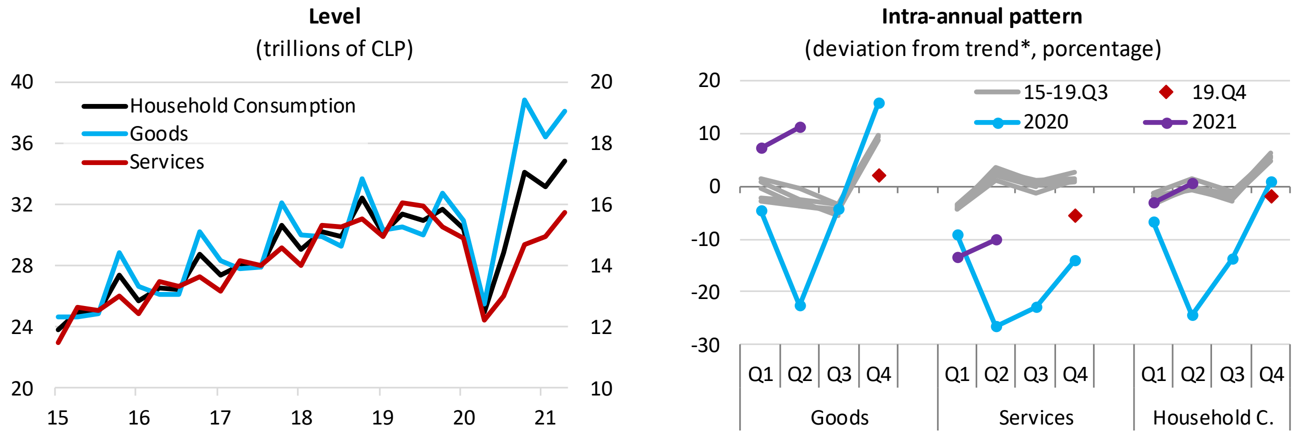
(1) Monthly indicator of economic activity
(2) Vertical lines mark the approval of withdrawals from pension funds
Source: Central Bank of Chile

2.1 Household consumption and payment data

Chilean Household consumption is compiled by the National Accounts Department at the Central Bank with a quarterly frequency and released 45 days after the respective quarter has ended, at the earliest. It is separated into services, durable and non-durable goods. They are constructed from VAT declarations and monthly surveys for retail sales and sales of durable goods and some other administrative data and surveys. They have traditionally not used payment data given the overlap with VAT declarations and the broader coverage of the latter.

As can be seen from figure 2, until 19.Q3 the series were fairly persistent, and the intra-annual patterns were quite predictable. 19.Q4 proved to be a shock to the system, as were 2020 and 2021. In this context, models that may have performed well until 2019 could be subject to large errors from then onwards. One would expect this to be the case for models that put a lot of weight on time series inertia, like SARIMA models. On the contrary, given the circumstances, models that lessen the weight of previous periods and focus on conjunctural signals will likely perform better. Enter payment data.

Figure 2: Household consumption



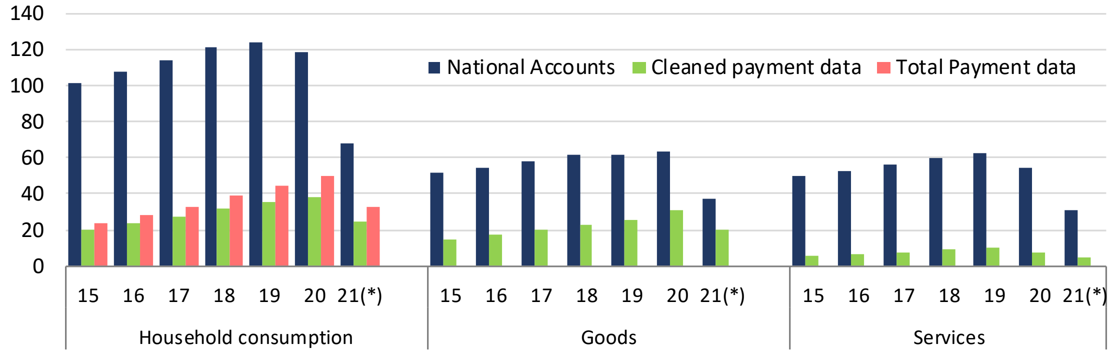
(*) Linear trend calculated over 15-19.Q3
Source: Central Bank of Chile

The payment data that I use is provided by Transbank. This company has for years been the sole provider of transaction intermediation. In the last couple of years others have appeared, but they still use Transbank's infrastructure. If you pay by card in Chile, debit or credit, it is extremely likely that Transbank will be managing your payment. For confidentiality reasons, the company does not provide the raw data to the Bank. The agreement with the provider is focused on acquiring data to evaluate macro issues, so it is comprised only of daily sales disaggregated by economic sector of the last link in the chain. This presents a challenge in that the same purchase could be classified as retail if it was acquired directly from the store or as a financial service if bought through a third party that is classified as such. However, given the level of detail of the information provided, there is nothing that can be done with it to account for such an issue.

For the purpose of nowcasting consumption, the payment data is filtered into three groups: goods, services and series to be excluded. The first two groups are self-explanatory and are defined according to the National Accounts definition. The excluded group contains all the sectors that seem too noisy to be useful. For example, those that start late in the sample or exhibit large sporadic leaps. One of these is auxiliary financial services. It shows an explosive growth at the very end of the sample that would be related, among other things, to the rapid growth of e-commerce services. The motivation for separating goods and services comes from the fact that payment data is biased towards goods as compared to consumption as measured by National Accounts. This can be seen from figure 3 where the value of payment data for goods increases from about a third to half of expenditure at the end of the sample, while for services, although it increases, the share remains relatively low.

Figure 3: Household consumption and payment data

(trillions of CLP)



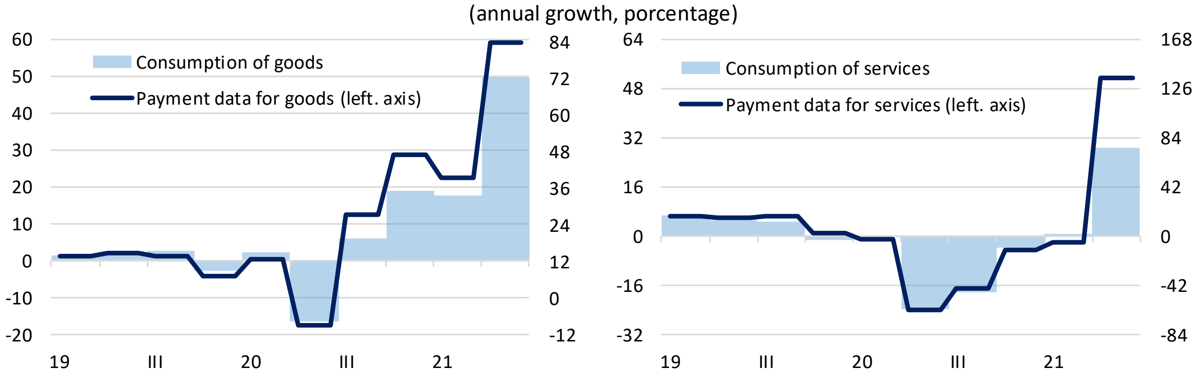
(*) First semester 2021

Source: Central Bank of Chile with data from Transbank

Although no adjustment is made to account for the fact that the composition of the payment data within these two large groups may differ from that of consumption, the yearly growth of both series is clearly related to its corresponding consumption line. Keeping in mind that the axes have been set so that the growth rates overlap visually, figure 4 shows them for both goods and services.

Figure 4: Household consumption and payment data

(annual growth, percentage)

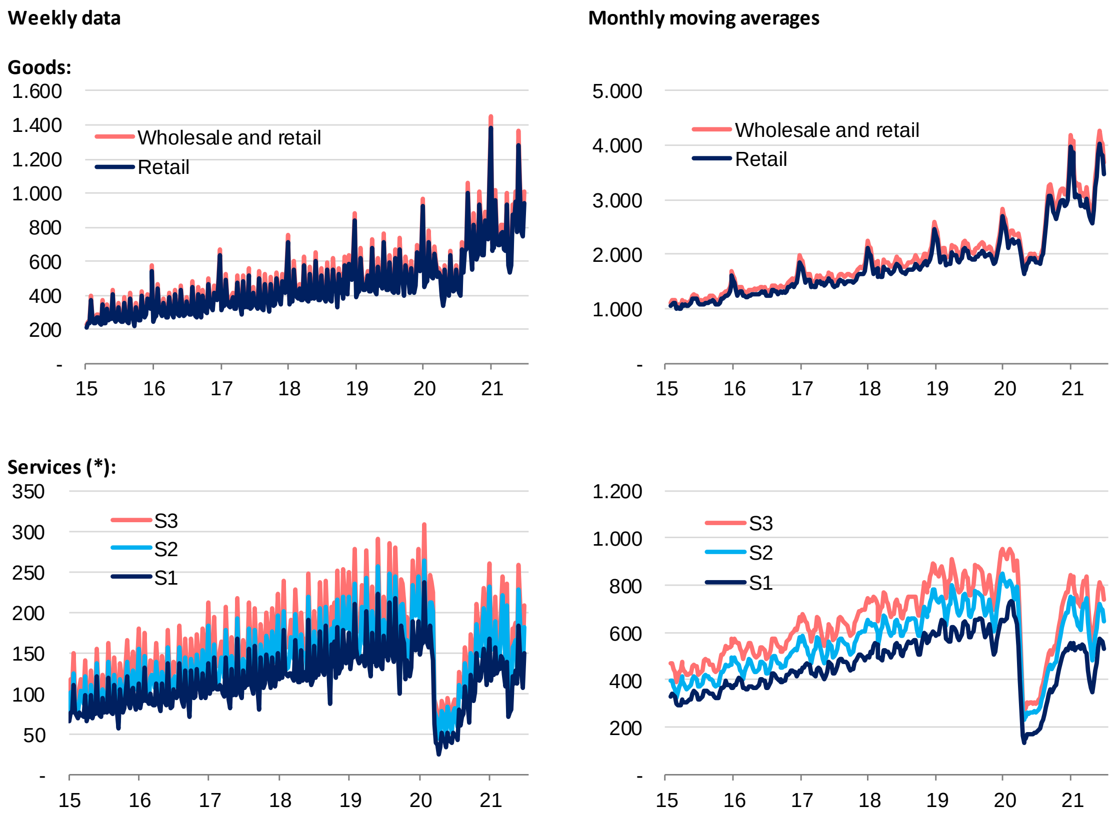


Source: Central Bank of Chile with data from Transbank

Looking at the data in this way already provides some insight. For goods it becomes obvious that something happens in the relationship between consumption and payment data in 20.Q3. The growth rates for payment data increase by a fair amount from then on. On the contrary, the relationship between both measures of services seems relatively stable throughout the whole sample except for the last period. This is probably a direct implication of the quarantines prevalent as of 20.Q2 and the income preserving measures, that is fiscal transfers and pension fund withdrawals. The additional liquidity of the households during this period could only be spent on goods, through e-commerce, not on services that were restricted. The deviation between measures for services in 21.Q2 could just be the extremely low levels from 20.Q2 affecting the annual growth rate.

As a it was mentioned, the "well-behaved" sectors were divided between those that provide goods and those that provide services. However, the separation is based on who is providing the good or service, not on who is consuming it. We could argue that purchases in wholesale stores may include some transactions that may not constitute household consumption, thereby introducing unwanted noise. Another issue may be that coverage of the payment data of one of the sectors may be completely unrepresentative of the sector as a whole. For this reason, two alternative measurements are used for goods and three for services. Following the previous argument for goods, there is retail alone and wholesale and retail. For services the smallest aggregate contains restaurants, private health, transport and personal services. A second aggregate adds private education and a third adds hotels and business services. The evolution and differences between them can be appreciated in figure 5.

Figure 5: Aggregations for goods and services
(billions of CLP)



S1: Restaurants + private health + transport + personal services.
S2: Private education + S1
S3: Hotels + Business services + S2

Source: Central Bank of Chile with data from Transbank

2.2 Modelling approach and specification

In the midst of the unusual conditions seen in the last two years and without a sufficiently long sample to even estimate the relationships in normal times, parsimony and risk reduction were paramount in the modelling approach. For this reason, a thick modelling approach was followed. The idea is motivated by Granger and Jeon (2004) and has to do with avoiding the choice of a single specification within a group of plausible models but rather extracting the useful information that each one of them may provide. Such a strategy seems sound, given that the persistence of accuracy of forecasting models may change abruptly (Aiolfi and Timmermann, 2006). This literature provides support and is closely related to forecast combination, the benefits of which are well documented (Aiolfi et al., 2011). Additional to the thick modelling, some simple refinements are included so as to account for some shocks that affect the payment data.

2.2.1 Bridge Models using payment data

The first step in this framework is to produce the nowcasts with the available data. Due to the short sample and the prevailing shocks, the implementation involves defining three simple bridge equations to nowcast all series. These make economic sense and account for the characteristics of the data in different ways so that hopefully they may not all fail simultaneously when hit by a shock.

The general equation is described by:

$$Y_t = \alpha + \sum_{i=1}^{N} \beta_i x_{i,t}^{Q} + \epsilon_t$$

where Y_t is National Accounts consumption in trimester t, $x_{i,t}^{Q}$ is the daily payment data i converted to quarterly frequency for trimester t with N being the number of regressors. The way in which the latter is done is described in the next section. The three specific equations are the following:

$$dlog(Y_t) = \gamma_0 + \gamma_1\, dlog(Y_{t-4}) + \gamma_2\, dlog(x_{i,t}^{Q}) + \gamma_3\, AR(1) + \varepsilon_t$$

$$dlog(Y_t) = \gamma_0 + \gamma_1\, dlog(x_{i,t}^{Q}) + \gamma_2\, AR(1) + \gamma_3\, AR(4) + \varepsilon_t$$

$$log(Y_t/Y_{t-4}) = \gamma_0 + \gamma_1\, log(x_{i,t}^{Q}/x_{i,t-4}^{Q}) + \gamma_2\, AR(1) + \varepsilon_t$$

where $AR(L)$ introduces an autoregressive structure for the error term including lag L of it and ε_t is the contemporaneous error.

The first six nowcasts for household consumption are produced using the three equations with the total payment data and the total cleaned payment data.

Proceeding in a similar fashion for goods and services produces six different nowcasts for goods and nine for services. By adding the nowcasts of every possible pair of forecasts, one for goods and one for services, produces 54 bottom-up forecasts for household consumption.

As an alternative to the bottom-up approach, 18 additional nowcasts are produced by including all possible pairs of goods and services in the same equation by substituting:

$$\gamma_*\, dlog(x_{i,t}^{Q}) = \gamma_{*,G}\, dlog(x_{G,i,t}^{Q}) + \gamma_{*,S}\, dlog(x_{S,i,t}^{Q})$$

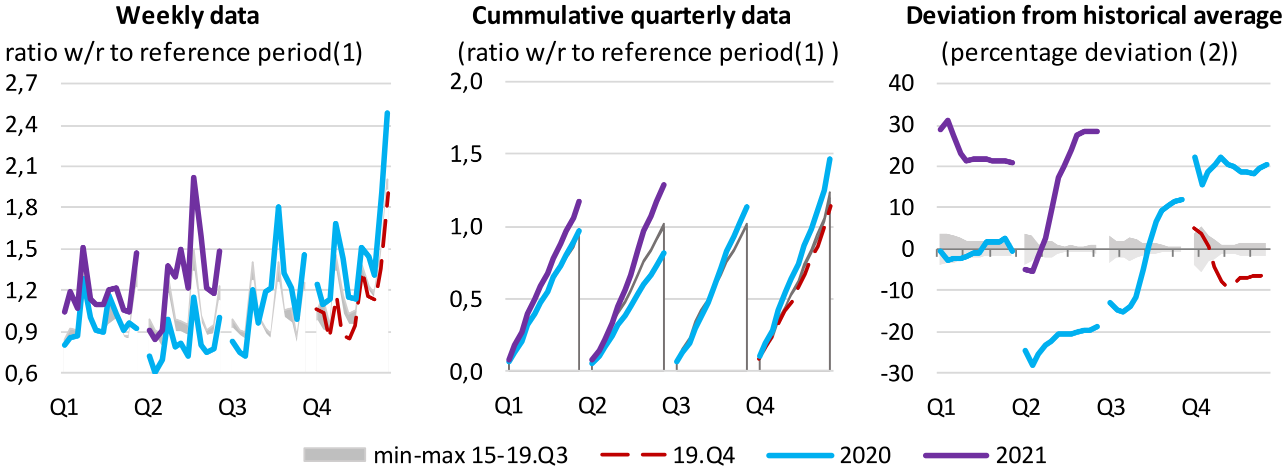
Using each one of these specifications results in 78 nowcasts for household consumption at any point in time.

2.2.2 Nowcast update with every new week of data

The previous section assumes that the payment data has been transformed to quarterly frequency. The way it was initially done involved dividing trimesters in 12 weeks with cutting points every 7th, 14th, 21st and end of the month. Then the data was accumulated up to the week number that was available for the last observation for the whole series. For example, if the last available observation was the 5th week of 20.Q1, the whole series that entered the regression was made up of trimesters that contained only the accumulation up to the 5th week of the respective period.

Following this procedure assumes that expenditure evolves in a fairly predictable manner and that patterns repeat themselves from one trimester to the next. To see whether such an assumption is reasonable, figure 6 presents different representations of the intra-annual regularity of expenditure based on the payment data. These measures present the ratio with respect to the average in the first quarter[1]. The left panel presents the weekly data, the middle panel the quarterly accumulation and the right panel presents the information in the middle panel in terms of deviation with respect to the average quarterly accumulation for the period before the social unrest episode. As can be seen from the grey areas in all graphs, the dispersion for the 15-19.Q3 period is relatively low and therefore the approach presented in the previous paragraph is probably reasonable.

Figure 6: Intra-annual behaviour of payment data



(1) corresponding reference period is the first quarter of each year. For 2020 y 2021 a "normal" first quarter is estmated based on 19.Q1 and average historical growth for that quarter. (2) Deviation w/r to average cummulative quarterly expenditure 2015 - 2019.Q3.
Source: Central Bank of Chile with data from Transbank

However, as of 19.Q4 sudden and significant decreases and increases in spending occur. This is most easily seen from looking at the right panel. One of the most dramatic examples is 20.Q3, when spending starts around 15% under the historical average jumping midway through the trimester to end up 10% above. Similarly, in 21.Q2 spending starts around 5% under the historical average jumping midway through the trimester to 30% above. It becomes obvious that during this period the assumption on which the simple weekly accumulation relies seems less appropriate.

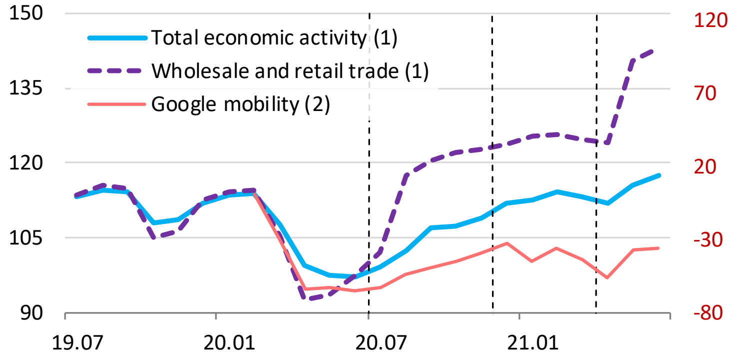
Although this conclusion is evident after the events, by 19.Q4 it would already have been clear that large shifts in spending could occur and would negatively affect the simple approach. More so, the gap that started to appear in June 2020 between retail trade and mobility, which can be appreciated in figure 7, and increased dramatically as of August, would have suggested a shift towards electronic payment methods due

[1] Because of the respective contraction and boost in expenditure in the first quarters of 2020 and 2021, the ratios are calculated over a counterfactual first trimester constructed using 2019.Q1 and the average historical growth rate.

to restrictions of movement. Something must have changed if mobility for retail and recreation remains 30% under normal times but spending in these sectors is 30% over.

Figure 7: Economic activity and mobility

(Index 2013=100, seasonally adjusted; deviation w/r ref. period)



(1) Monthly indicator of economic activity

(2) Google mobility retail and recreation. Reference period: february 2020.

(3) Vertical lines mark the approval of withdrawals from pension funds

Source: Central Bank of Chile and Google

The way in which these two issues are dealt with is somewhat different. As regards the sudden shifts in spending, as can be observed in the weekly data in figure 6, the intra-quarterly dynamics do not change very much once the shifts are accounted for. This has to do with the fact that spending tends to cluster around both sides of the first day of the month and certain holidays. Then, instead of discarding the information of the whole trimester by introducing dummy variables every time this occurs, the shifts are distributed proportionally over the whole trimester by using the average historical intra-quarterly pattern and the spending of the whole trimester. Then, as an example, if the last available observation is the 5th week of 20.Q1, the whole series that enters the regression is made up of trimesters equal to the spending that would have been made up to the 5th week if the actual spending in each trimester from 19.Q4 and before had followed the historical distribution for that trimester. Obviously, this approach does not help to predict sudden shifts in expenditure, but it eliminates the noise they produce in the estimation of coefficients, without losing any observations. As regards the apparent relative increase in the use of electronic payment methods as of 20.Q3, appropriate dummy variables are included in the bridge equations.

2.2.3 Obtaining the final nowcast

The aforementioned procedure produces 78 individual nowcasts for household consumption at any point in time. In this context and building on the forecast combination literature, two alternative point-estimates are produced. The first is the median nowcast and the second is the weighted average of nowcasts where the weights are determined based on the root mean square error (RMSE) of the yearly growth rate over the previous four trimesters. In such a situation, often the inverse of the RMSE is used directly as the weighting factor. However, as has been well documented, when the number of competing models is moderately large, such a strategy tends to produce weights that do not differ much from equal weights (Aiolfi et al., 2011). To have an alternative to the median nowcast, I use a weighting factor for nowcast *j* at time *t* given by the following formula:

$$f_{j,t} = \left(100\left[IR_{j,t} - min(IR_t)\right]\right)^4 \quad where \quad IR_{j,t} = \frac{1}{RMSE_{j,t}} \quad and \quad IR_t = \{IR_{1,t}, \dots, IR_{J,t}\}$$

The idea behind this formula is to exponentially broaden the distance between the better and worse performing models.

3. Pseudo real-time nowcasting exercise

3.1 Design of the empirical analysis

The exercise is quite simple. I use 2015 to 2017 to start estimating the coefficients and produce the nowcasts for 18.Q1 with every new week of data. After the trimester is over, I incorporate 18.Q1 into the estimation of coefficients and proceed with the next trimester. As of 18.Q4 I start calculating the median nowcast and the nowcast based on the previous period's out-of-sample forecasting error all the way to 21.Q2. I then calculate the RMSE over two periods: 18.Q4 to 19.Q3, or pre-crisis period, and 19.Q4 to 21.Q2 for the crisis period. These results are then compared to those of the benchmark models.

3.2 Benchmark models

3.2.1 SARIMA

SARIMA models are used widely as benchmarks in the forecasting literature and therefore make an ideal choice for this exercise. Robust automatic routines have been developed and in this case I turn to the one that comes in Eviews. In short, it chooses the integration order of the series following a routine that makes use of the KPSS test (Kwiatkowski *et. al.*, 1992) and then the lag order according to the Bayesian Information Criterion:

$$BIC = \ln\left(\frac{1}{T}\sum_{i=1}^{T} e_i^2\right) + \frac{k}{T}\ln(T)$$

where $\sum_{i=1}^{T} e_i^2$ is the sum of square residuals, k the number of parameters and T the number of observations.

3.2.2 Traditional bridge models

Although SARIMA models have a great track record in stable times, given that the period under scrutiny includes large shocks, models that include conjunctural information will probably perform better. Prior to having access to the payment data, bridge models based on monthly data were already in use at the Bank. These are a more obvious benchmark against which to compare the payment data models. The forecasting approach and the two original equations are described in Cobb et al. (2011) and were specified balancing the economic sense of the relationships with the availability of data. One of the equations adheres strictly to the use of variables that are directly related to consumption while the second includes others that empirically seemed to have some predictive power. The approach uses the actual data that would have been available for the monthly indicators at the time of forecasting and fills in the rest of the trimester with SARIMA forecasts produced using the same Eviews routine presented in the previous section. This preserves the real-time forecast evaluation approach of the exercise. Acknowledging that over time these relations may change, two additional equations are used based on the same variables as the originals, except that the lag orders are defined using Stepwise-Forwards Regression. The monthly conjunctural variables that are used in these equations are: waged employment, consumption imports, supermarket sales, consumer confidence, real value of the stock exchange, retail sales, real wages, interest rate, real exchange rate and car sales. The hypothetical cutting date for the availability of actual data is set at the 15th of every month. This means that all the data that is released at the end of the current month or in the few days following is considered as available for the models. Data that is closer to being available with a one-month lag is considered not to be available. The cutting date is also coherent with having the payment data for that month.

3.3 How does the exercise differ from a true real-time experiment

There are two reason why this exercise is only pseudo and not truly real-time. On the one hand it is not accounting for revisions in the data and, on the other hand, it does not reflect that household consumption of the previous trimester is released with a minimum of 45 days after the trimester has ended and not immediately. This is relevant because the nowcasting models specified in quarterly differences pivot their current period nowcast on the previous trimester. However, for reasons that are presented in the following paragraphs, neither of these problems is considered to be too serious in this particular case.

The problem of data revision has to do with the fact that data that is available at the time of forecasting may differ from the definitive or final versions, in some cases quite substantially. For example, GDP is revised many years after it is first published. This, because the initial estimate is based on partial information that is then completed as time goes by. Other conjunctural indicators also suffer modifications for this same reason. This may be a problem in a forecasting context for multiple reasons. One is that the variable of interest may be the final version but due to the publication lag it is not possible to establish the relative performance of different models within a time frame that makes the results relevant. Another is that if it is the independent data that suffers revisions, using the revised version in the accuracy comparison may overstate the real-time predictive content of some variables (Stark and Croushore, 2002).

From the perspective of conjunctural analysis, probably the most relevant release is the first because it is the signal to which the economic agents react and therefore, for this exercise, the objective would be to predict the first releases of household consumption. Bearing this in mind, I consider the issue of data revision to be relatively minor. On the one hand, payment data is not revised and, on the other hand, the size of revisions to household consumption have been negligible when compared to the size of the shocks that have affected it in the last three years. The same is true for the data that undergoes revisions that are included in the benchmark models. The issue of the publication lag of the pivot trimester is also considered to be relatively unimportant for this analysis for two reasons. First, the apparent oversight affects both the payment data and benchmark models so the comparison between them is still fair. Second, and more important, is that the relevance of the nowcasting models increases with the passage of time. Then, when only a couple of weeks of the concurrent trimester have gone by it makes more sense to stick with the forecasting models and wait for more information to become available. Given the current publication lag, the problem disappears from the analysis from the 5th week onwards.

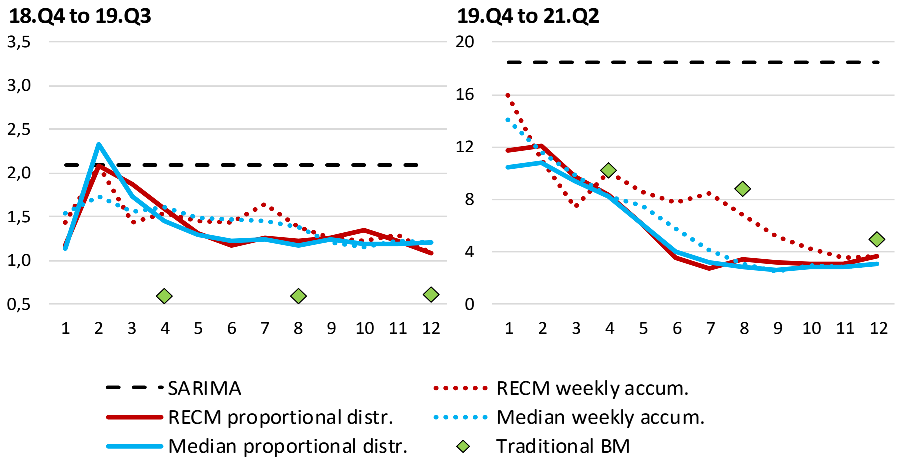
3.4 Out-of-sample nowcasting results

These are the main findings that are presented in more detail in the rest of this section. The results suggest that accuracy of the payment data models tends to increase with the accumulation of more data and most of the gain would be achieved mid-way through the trimester. Both these things are true for the whole sample, as is the fact that the payment data models beat the SARIMA models by a fair amount. As regards the comparison with the traditional bridge models, the gains only become apparent in the turbulent period and the advantage would seem to disappear as the information for the whole trimester becomes available. This is not surprising given that the information included in the traditional models is used to construct the actual indicator. Another finding is that although the dispersion in the forecasts increases significantly in the turbulent period, the median and RECM weighted average continue to perform relatively well. When looking at the components, the dispersion of the nowcasts of goods becomes very narrow by the 12th week while for services in some periods it remains quite wide. This is probably motivated by the fact that the coverage of goods is much higher than for services. Finally, in line with the findings of Aiolfi and Timmermann (2006) the relative accuracy of the different specifications varies greatly from one period to the next, therefore warranting the use of the thick modelling approach. As a note of caution, it is worth remembering that all of these results are suggestive because of the limited amount of data on which they rely on.

3.4.1 Relative accuracy of the point-forecasts and their improvement with increasing data

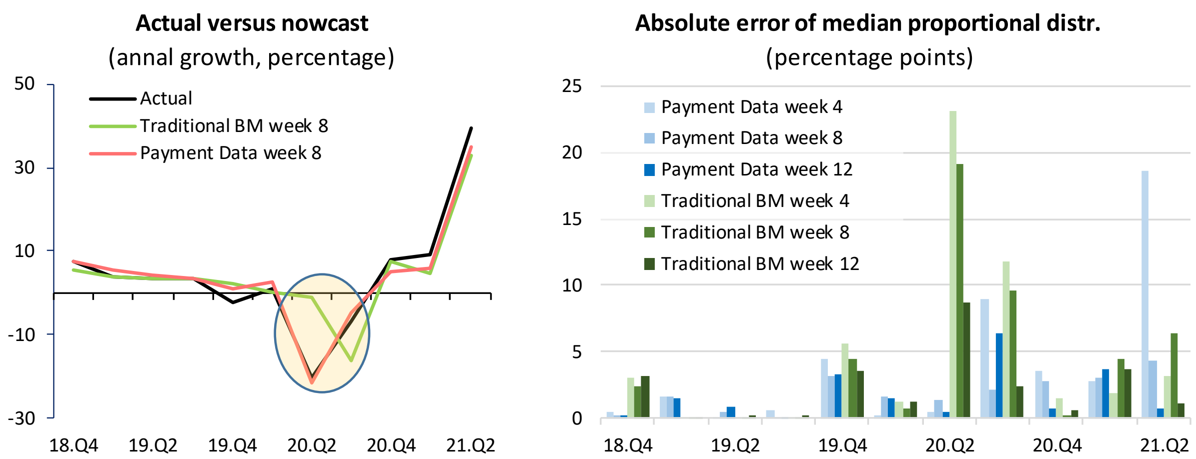
Figure 8 summarizes the main results of this paper. One of these is that in both normal and turbulent times all the measures for point-estimates improve as new data is incorporated. This is most obvious in the cases that use the proportional distribution of data which also goes to show that the approach does work as a way of evening out the sudden shifts in spending. The figure also shows that accuracy improves dramatically over the first couple of weeks and, in line with the arguments presented earlier in the paper and in light of the results for the traditional bridge models, it only makes sense to start considering the nowcasts from the payment data models from mid-trimester onwards. As regards the performance relative to the benchmark models, the payment data models beat the SARIMAs by a fair amount. In the normal period it roughly halves its RMSE from about the sixth week onwards while in the turbulent period, due to the predictably appalling performance of the SARIMAs, RMSE of the payment data models is about five times smaller.

Figure 8: RMSE of the annual growth of nowcasting models as weekly data accumulates



As regards the comparison with the traditional bridge models, the gains only become apparent in the turbulent period and the advantage would seem to disappear as the information for the whole trimester becomes available. The superiority of the traditional BMs in normal periods is probably due to the high persistence of the series and the broader coverage of their monthly indicators. The relatively good performance of the SARIMA models supports the first reason, but it cannot be the only one because the traditional BMs are also far superior to the SARIMAs.

Figure 9: Breakdown of RMSE

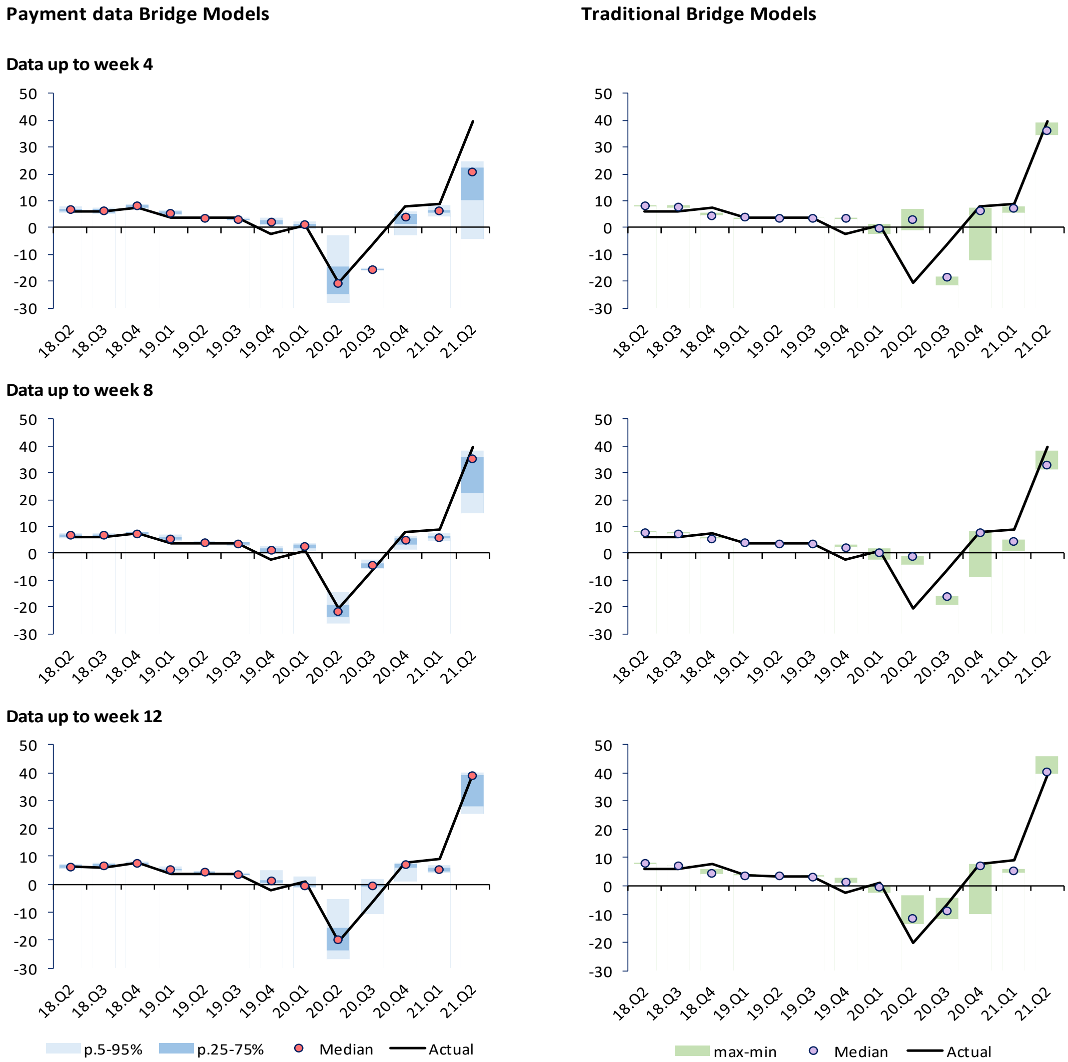


On the contrary, in turbulent times the payment data models, on average, pick up the changes in consumption far earlier than the traditional models. This would suggest that the persistence that is built into the latter and that potentially gives them the edge in normal times plays against them when a shock hits. Going into this point in a bit more detail, figure 9 shows the forecasts at week 8 in contrast with the actual data and the absolute error for the models using data up to week 4, 8 and 12. As can be seen from the left panel it would seem that there are only two periods in which the traditional models are completely lost: Q2 and Q3 from 2020. In the usual high-persistence time-series fashion, they completely overshoot the contraction of Q2 and by incorporating the last period's error they miss the relative recovery of Q3. From looking at the absolute errors, it becomes clear that these two periods are the offending events to the otherwise reasonable performances of the traditional models. Another thing that can be appreciated is that even within the turbulent period, besides the aforementioned periods, it is not evident whether the payment data models are superior. For the social unrest in 19.Q4, for example, the errors are quite similar at any level of accumulation of data.

3.4.2 Relative accuracy of the set of models

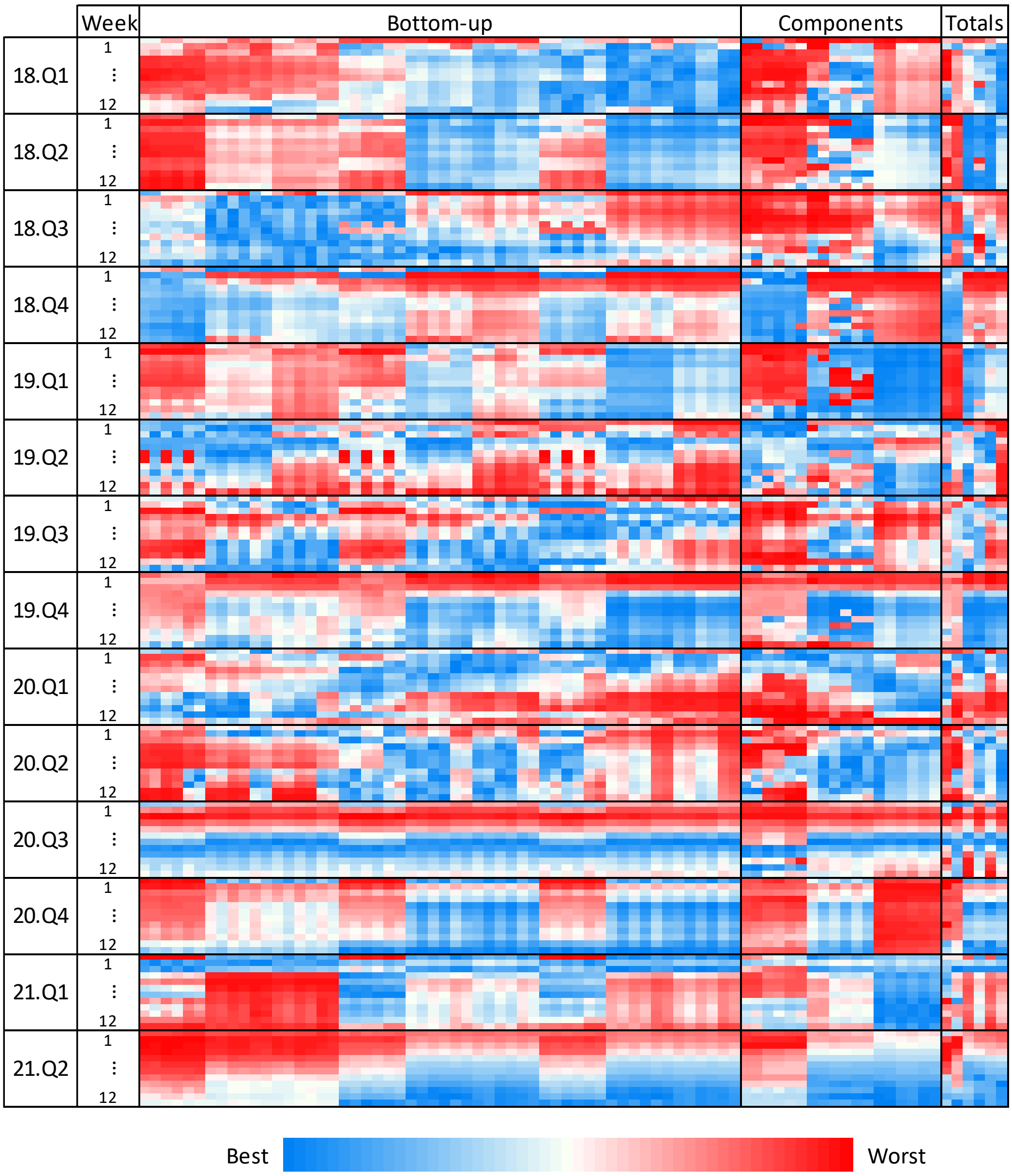
All the previous analysis is conducted in reference to the nowcast provided by the median and weighted average. It is a fair question to ask whether this is the best approach in this case. Figure 10 presents the dispersion of nowcasts for both the payment data and traditional bridge models. A quick look makes it clear that, even though it improves as more data is available, the dispersion bands remain quite wide in some periods even when all the information is available. To understand why this happens it is necessary to look in more detail into the consistency of the performance of the individual models. If one or a reduced number of models are really good and all of the others are just noisy it would make more sense to drop the latter completely and only use the well performing ones.

Figure 10: Dispersion of nowcasts over the evaluation period at weeks 4, 8 and 12 (annual growth)



To look at this point, figure 11 presents the rank in accuracy for every trimester of each individual model for each week of accumulated data. That is for each trimester, a model *j* using the data accumulated up to week *w* is given a rank number going from 1 to 936 (78 individual models * 12 possible levels of accumulated weekly data) depending on its nowcast absolute error. For a model *j* improvements due to data accumulation are reflected by a change in colour that goes from “closer to red” to “closer to deep blue” as one descends. This can be seen in many of the models in 21.Q2 and some other periods but it is not obvious in the entire sample. The predominance of a model *j* over all other models would be reflected by a “closer to deep blue” column that spans the whole figure. Such a column cannot be appreciated in figure 11. In fact, the opposite is true. The clusters of deep blue skip horizontally with nearly every change of trimester and in many cases from one week of accumulated data to the next.

Figure 11: Quarterly relative accuracy persistence of models
(rank)

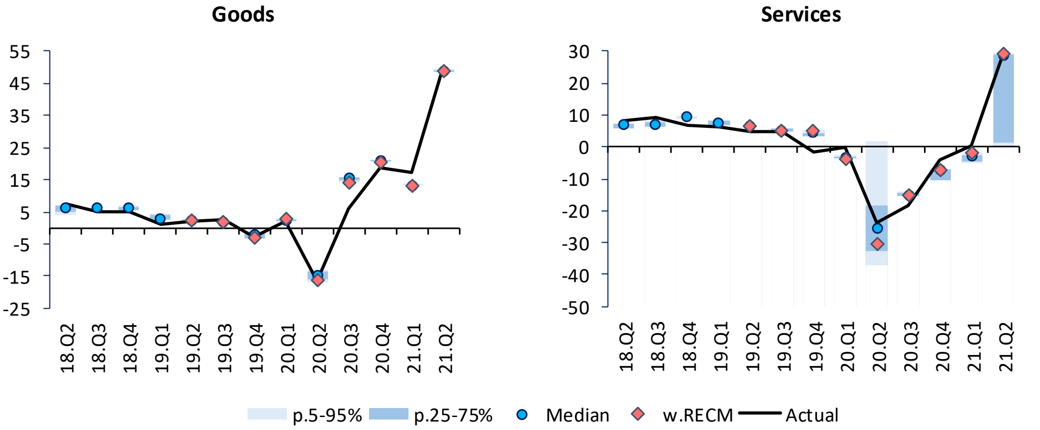


Bottom-up: 54 nowcasts as the sum of every possible pair of individual forecasts for goods and services,
Components: 18 nowcasts produced by including all possible pairs of goods and services in the same equation.
Totals: 6 nowcasts using the three equations both with the total payment data and total cleaned payment data.

This behaviour of the performance of the individual models suggests that the difficulty in identifying the best model stems from the fact that one does not exist. This goes to support the findings by Aiolfi and Timmermann (2006) regarding rapidly changing relative performance of forecasting models and might explain why the weighted average point-estimate does not outperform the comparatively simple median.

As regards explaining the source of the uncertainty surrounding the forecasts, despite its being only a subset of the nowcasts for household consumption, the set produced independently for goods and services presented in figure 12 goes on to suggest that it is the services component of the nowcasts that contribute with most of the variability,[2] as one might expect from the reduced coverage when compared to the National Accounts data.

Figure 12: Dispersion of forecasts over the evaluation period for goods and services at week 12 (annual growth)



4. Concluding remarks

This paper shows the results of using electronic payment data to track Chilean household consumption in real time during the pandemic and compares the results to those of standard nowcasting methods used at the Central Bank. The results show that the simple models that rely solely on payment data only beat the competing traditional bridge models as the shocks occur. The gain is not obvious in more stable times and once the shocks have been incorporated into the estimation of the benchmark models. Also, in developing the payment data models it became clear that this source of naturally occurring data is being affected by two distinct idiosyncratic shocks. On the one hand, there are sudden surges or contractions in spending that require accounting for in the modelling framework. On the other hand, there is at least one time in which there is a relative increase in electronic payment usage that unattended would exaggerate the nowcast for household consumption. Both these findings, in addition to the fact that a large share of transactions are not made through electronic payment, suggest that the availability of high frequency data does not undermine using more traditional methods that rely on less frequent data. It makes more sense to use this new scope of models in conjunction with other proven methods and judgement.

Regarding this last point and reflecting on the strengths of the payment data models and the traditional bridge models, there is a strong argument towards using judgment to assign the subjective weights to each type of model instead of blindly combining them. This, because the defining characteristic of nowcasting is that the period of interest is occurring concurrently with the prediction and it is therefore possible to know whether the period being nowcasted is being affected by a shock or not. In this line, it also makes sense to

[2] The Annex presents many of the figures that are presented in the body of the text but for goods and services.

make certain that in developing the payment data models further they do not lose their relatively better qualities for reflecting conjunctural shocks.

In light of this and as regards ~~to~~ further research, at least two avenues come to mind. The most obvious is, as more data becomes available, to develop models that incorporate both this daily high frequency payment data and monthly indicators but that retain the strengths of both sources of data. One would expect that the interaction of these two sources of information could result in gains in accuracy. Another line of research has to do with making a greater effort to incorporate the actual National Accounts weights into the payment data that goes into the models. In the current approach the difference in coverage is accounted for only at the level of goods and services but not within these groups, so some of the subgroups may be over/under-represented in the payment data. Assigning each economic sector that is in the payment data with the appropriate representativity could lead to further gains in nowcasting accuracy.

References

Aiolfi, M., Capistrán, C., Timmermann, A. (2011). “Forecast combinations”. In. The Oxford Handbook of Economic Forecasting. Ed. By M. Clements and Hendry, pp. 355-388.

Aiolfi, M., and Timmermann, A. (2006). "Persistence in Forecasting Performance and Conditional Combination Strategies." Journal of Econometrics, 135(1-2), 31-53.

Aastveit, K., Fastbø, T., Granziera, E., Kenneth Sætherhagen Paulsen, K. and Næss Torstensen, K. (2020). “Nowcasting Norwegian household consumption with debit card transaction data”, Norges Bank’s Working Papers, Working Paper 17/2020

Cobb, M., Echavarría, G., Filippi, P., García, M., Godoy, C., Gonzalez W., Medel, C. y Urrutia, M. (2011). “Short-Term GDP Forecasting Using Bridge Models: a Case for Chile”, Documentos de Trabajo N° 626, Banco Central de Chile, Mayo 2011.

Conesa, C., Gambacorta, L., Gorjon, S. and Lombardi, M. (2015) “The use of payment systems data as early indicators of economic activity”, Applied Economics Letters, 22:8, 646-650

Galbraith, J.W. and G. Tkacz (2018), “Nowcasting with payments system data”. International Journal of Forecasting, 34, pp. 366-376.

Gil, M., Pérez, J. and Urtasun, A.(2019), "Nowcasting private consumption: traditional indicators, uncertainty measures, credit cards and some internet data," IFC Bulletins chapters, in: Bank for International Settlements (ed.), The use of big data analytics and artificial intelligence in central banking, volume 50, Bank for International Settlements.

Granger, C. W. J. and Jeon, Y., (2004), "Thick modeling," Economic Modelling, Elsevier, vol. 21(2), pages 323-343, March.

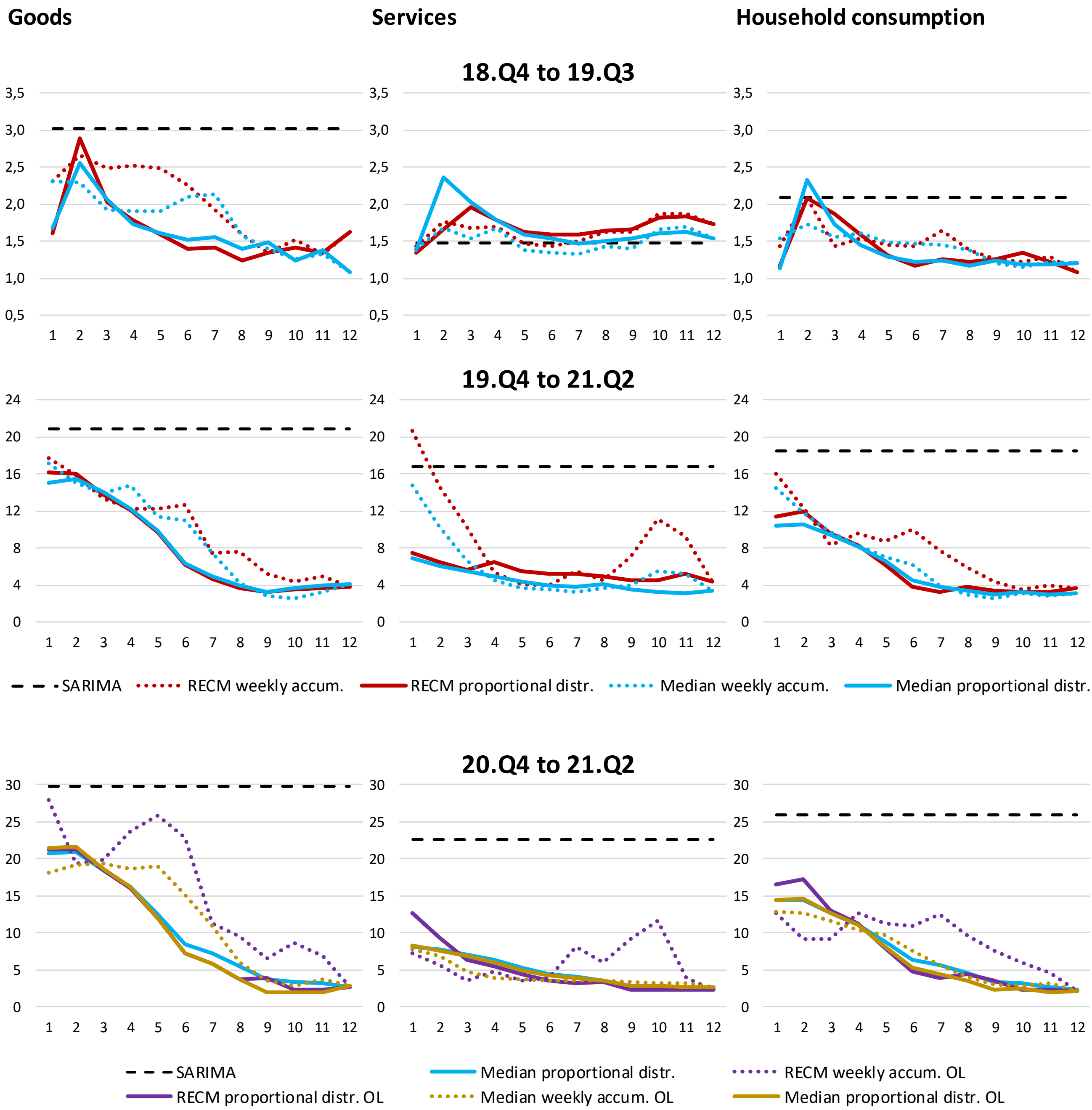
Guerrieri, V., G. Lorenzoni, L. Straub, and I. Werning (2020). “Macroeconomic Implications of COVID-19: Can Negative Supply Shocks Cause Demand Shortages? NBER Working Papers 26918, National Bureau of Economic Research, Inc.

Kwiatkowski, D.; Phillips, P. C. B.; Schmidt, P.; Shin, Y. (1992). "Testing the null hypothesis of stationarity against the alternative of a unit root". Journal of Econometrics. 54 (1–3): 159–178.

Stark, T y Croushore, D. (2002). “Forecasting with a Real-Time Data Set for Macroeconomists”. Journal of Macroeconomics 24 (December): 507–31

Annex

Figure A.1: RMSE of nowcasting models as weekly data accumulates



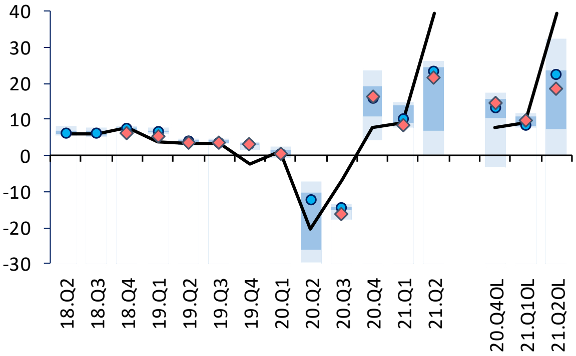
Note: OL denotes the models including the dummy variable to account for the relative increase in electronic transactions.

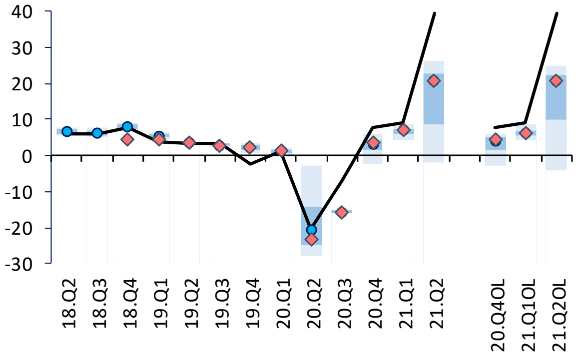
Figure A.2: Dispersion of forecasts over the evaluation period at weeks 4, 8 and 12

Household consumption

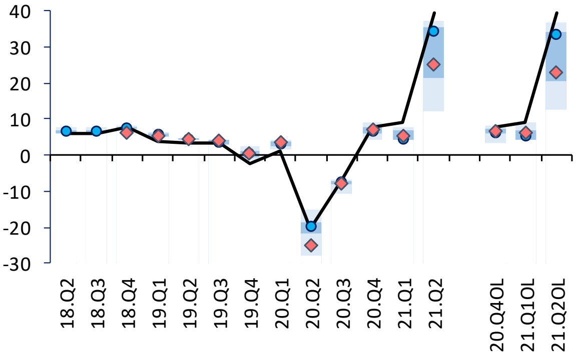
Weekly acummulation | **Proportional distribution**

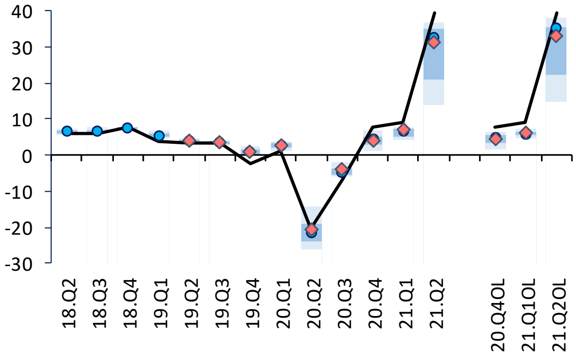
Data up to week 4



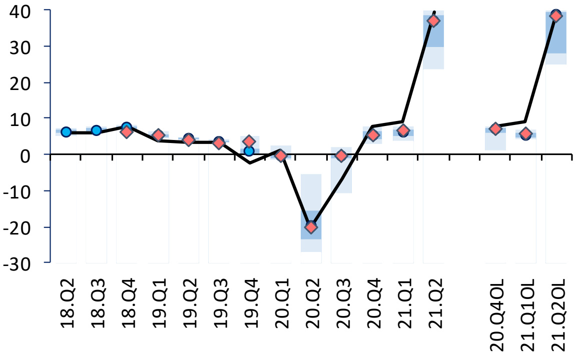


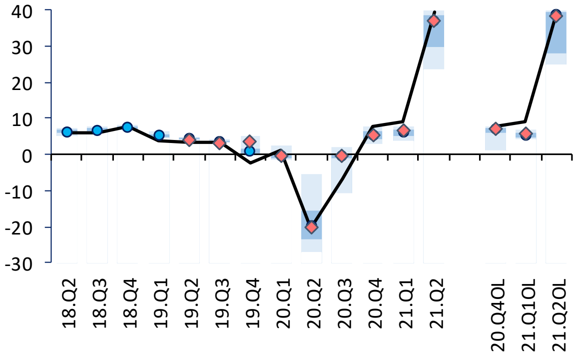
Data up to week 8





Data up to week 12





p.5-95% p.25-75% Median w.RECM Actual

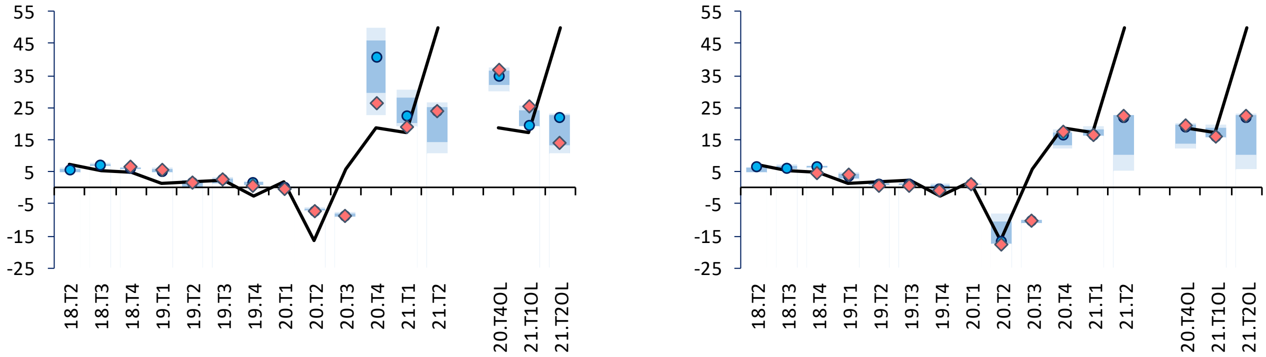
Note: OL denotes the models including the dummy variable to account for the relative increase in electronic transactions.

Figure A.3: Dispersion of forecasts over the evaluation period at weeks 4, 8 and 12

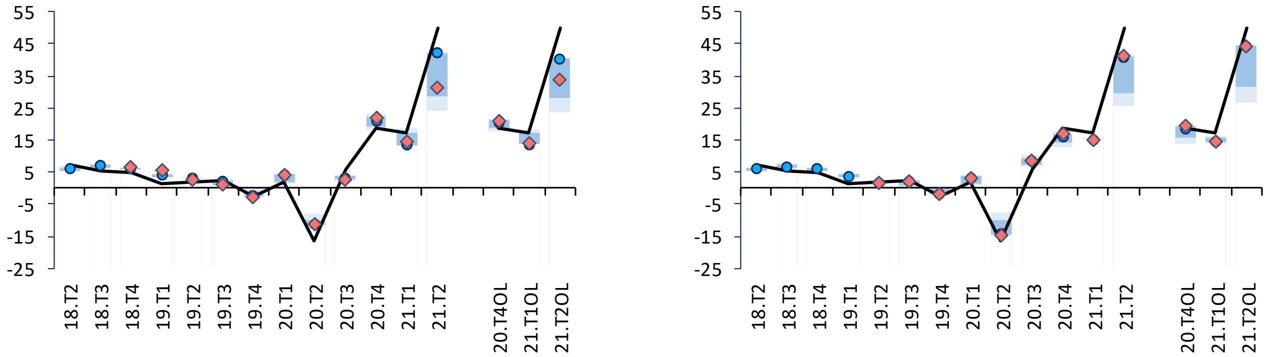
Consumption of Goods

Weekly acummulation | **Proportional distribution**

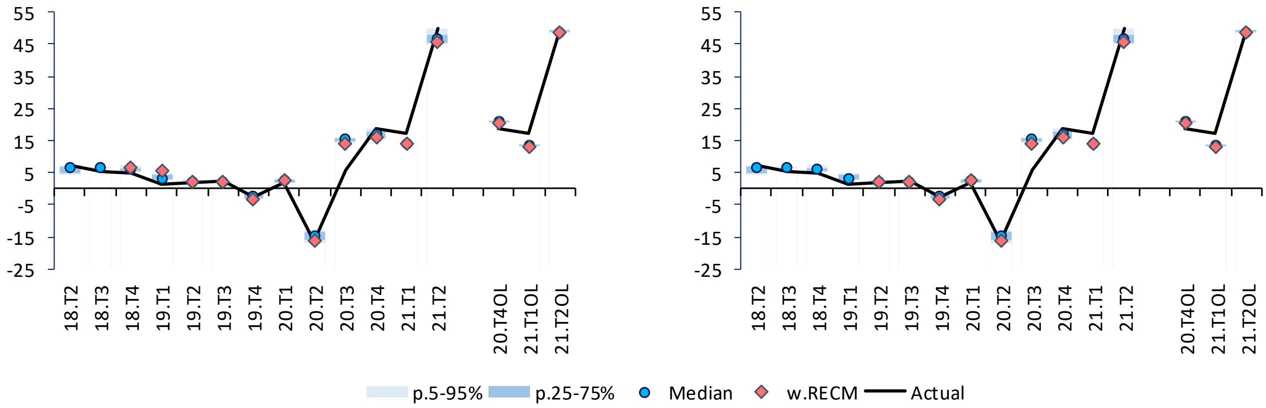
Data up to week 4



Data up to week 8



Data up to week 12



Note: OL denotes the models including the dummy variable to account for the relative increase in electronic transactions.

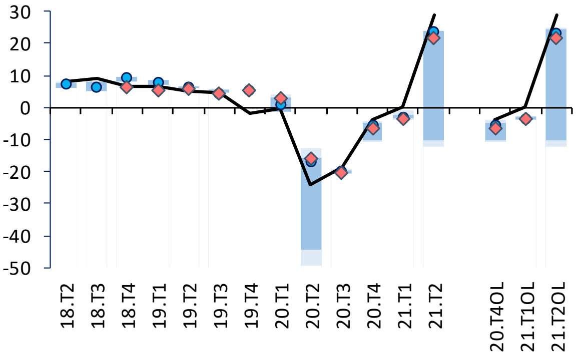
Figure A.4: Dispersion of forecasts over the evaluation period at weeks 4, 8 and 12

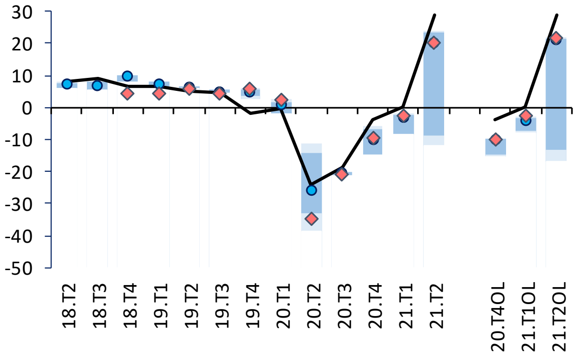
Consumption of services

Weekly acummulation

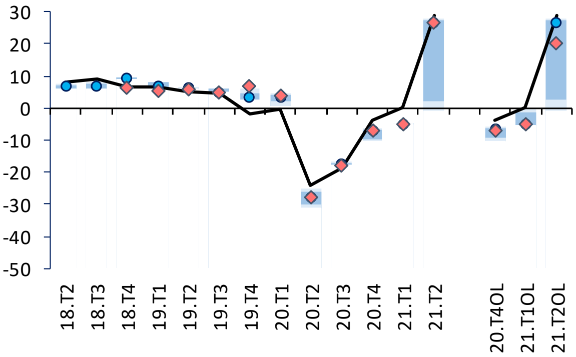
Proportional distribution

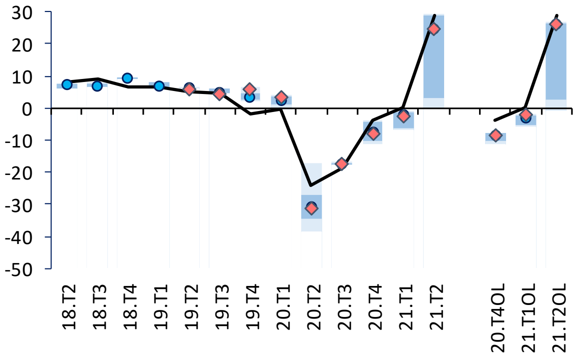
Data up to week 4



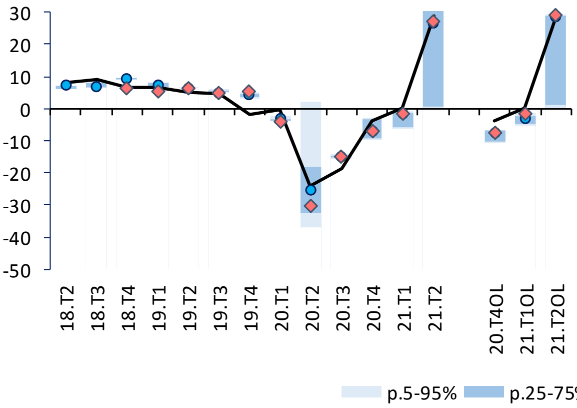


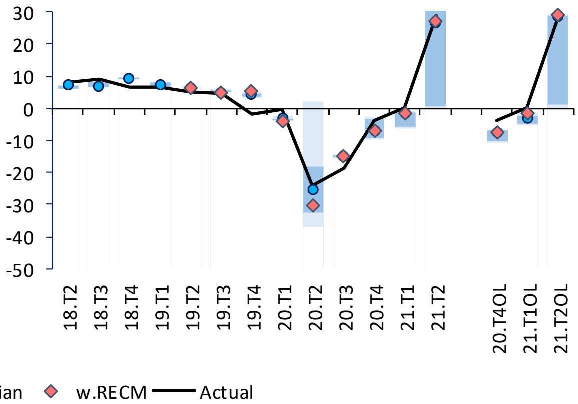
Data up to week 8





Data up to week 12





Note: OL denotes the models including the dummy variable to account for the relative increase in electronic transactions.

Documentos de Trabajo **Banco Central de Chile**	**Working Papers** **Central Bank of Chile**
NÚMEROS ANTERIORES	PAST ISSUES
La serie de Documentos de Trabajo en versión PDF puede obtenerse gratis en la dirección electrónica: www.bcentral.cl/esp/estpub/estudios/dtbc. Existe la posibilidad de solicitar una copia impresa con un costo de Ch$500 si es dentro de Chile y US$12 si es fuera de Chile. Las solicitudes se pueden hacer por fax: +56 2 26702231 o a través del correo electrónico: bcch@bcentral.cl.	Working Papers in PDF format can be downloaded free of charge from: www.bcentral.cl/eng/stdpub/studies/workingpaper. Printed versions can be ordered individually for US$12 per copy (for order inside Chile the charge is Ch$500.) Orders can be placed by fax: +56 2 26702231 or by email: bcch@bcentral.cl.

DTBC – 931
Nowcasting Chilean household consumption with electronic payment data
Marcus P. A. Cobb

DTBC – 930
Sentimiento en el Informe de Estabilidad Financiera del Banco Central de Chile
J. Sebastián Becerra, Alejandra Cruces

DTBC – 929
Inequality, Nominal Rigidities, and Aggregate Demand
Sebastian Diz, Mario Giarda, and Damián Romero

DTBC – 928
Rational Sentiments and Financial Frictions
Paymon Khorrami, Fernando Mendo

DTBC – 927
Saving Constraints, Inequality, and the Credit Market Response to Fiscal Stimulus
Jorge Miranda-Pinto, Daniel Murphy, Kieran James Walsh, Eric R. Young

DTBC – 926
Liquidez del sector corporativo chileno: estimación de disponibilidad de caja bajo escenarios de estrés
Jorge Fernández, Fernando Pino, María Ignacia Valencia

DTBC – 925
Procyclical fiscal policy and asset market incompleteness
Andrés Fernández, Daniel Guzman, Ruy E. Lama, Carlos A. Vegh

DTBC – 924
Reserve Accumulation and Capital Flows: Theory and Evidence from Non-Advanced Economies
Juan Pablo Ugarte

DTBC – 923
Estimates of the US Shadow-Rate
Rodrigo Alfaro, Marco Piña

DTBC – 922
A Semi-Structural Model with Banking Sector for Stress Testing Scenario Design
J. Sebastián Becerra, José Carreño, Juan Francisco Martínez

DTBC – 921
Monetary Policy Surprises in Chile: Measurement & Real Effects
Boragan Aruoba, Andrés Fernández, Daniel Guzmán, Ernesto Pastén, Felipe Saffie

DTBC – 920
The Effects of Firms' Lobbying on Resource Misallocation
Federico Huneeus, In Song Kim

DTBC – 919
Learning Your Own Ability
Carlos Madeira

DTBC – 918
Commodities Fundamental Model
Francisco Arroyo Marioli, Fernando Letelier

DTBC – 917
Firm Sorting, College Major, and the Gender Earnings Gap
Federico Huneeus, Conrad Miller, Christopher Neilson, Seth Zimmerman

DTBC – 916
Optimal Monetary Policy and Incomplete Information: Does the Real Exchange Matter?
Rodrigo Caputo, Felipe Leal

DOCUMENTOS DE TRABAJO Noviembre 2021