Forecasting Daily Demand in Cash Supply Chains

Michael Wagner
Department of Accounting, Hanken School of Economics,
Arkadiankatu 22, Helsinki, 00101, Finland

Abstract: Problem statement: Previous studies focused on explaining the long run determinants of currency demand offering limited insight into the short-run determinants and co-variability of daily demand in cash supply chains. Approach: This study contrasted competing techniques of forecasting daily demand in cash supply chains in order to determine the overall performance and the potential of joint forecasting for integrated planning. A joint forecasting approach was compared with well-established causal forecasting techniques, namely, a vector time series model and a seasonal ARIMA model using simple methods as benchmarks. Evaluation was based on multiple time series obtained from mid-size European bank with forecasting horizons of up to 28 days. Forecasting accuracy was measured using the mean absolute percentage error. Results: The seasonal ARIMA model resulted in a higher forecasting accuracy compared to the vector time series model. Variability in demand was mainly attributed to the day-of-the-week effect. Co-variability is captured by seasonality and calendar effects limiting the potential of joint forecasting. Cumulative forecasts for periods of 14 days are very robust with mean percentage errors of approximately two percent. Conclusion: The results confirmed the benefit of advanced forecasting techniques for daily forecasts. However, the study suggested that the role of information sharing is limited to coordination of replenishments across the cash supply chain and does not yield more accurate forecasts based on joint forecasting.

Key words: Cash supply chain, cash demand, forecasting, seasonal ARIMA, vector time series models

INTRODUCTION

Cash supply chains play a vital role for the efficiency of banking systems. This complex string of interrelations required to circulate currency from central banks to end users and back, is a special case of closed-loop supply chains. Currency circulates freely between different stages, such as central bank, Cash In Transit (CIT) providers, banks, Automated Teller Machines (ATM), corporate and private customers, incurring transaction and holding costs. Cash management decisions at each stage are interrelated, but mostly beyond the control of the central bank. Hence, currency in cash supply chains represents one of the largest autonomous liquidity factors in banking systems.

This study considers the co-variability of daily cash demand series within cash supply chains in addition to seasonality and calendar effects. The goal is to evaluate the overall forecasting accuracy of daily cash demand and the potential of joint forecasting in cash supply chains.

An analysis of forecasting daily currency demand in cash supply chains is interesting and important for mainly three reasons. From an economic perspective, it enhances the understanding of the demand for currency by considering seasonality and calendar effects in higher frequency data. Traditionally, data availability restricts studies to lower frequency data, such as monthly, quarterly and annual aggregates. From a supply chain perspective, it provides empirical evidence on the accuracy of forecasting demand in cash supply chains, which is essential for integrating autonomous cash management decisions across the cash supply chain, coordinating currency needs and increasing efficiency of banking systems. Similarly, from a business strategy perspective, the study provides further insights into the role of information sharing. In particular, the present study accounts for partnering, outsourcing and privatization that are currently transforming cash supply chains into a mature and optimized businesses (Rajamani et al., 2006).

Forecasting demand in cash supply chains requires an underlying theoretical framework that can reasonably explain the demand for currency. On theoretical grounds, the nature and structure of the demand for cash has been intensely debated by academics in the past (Serletis, 2007). Research has focused on the long run determinants of currency demand. Topics include estimation of income and interest rate elasticity, structural changes from currency
change over, influence of alternative payment systems, money hoarding and informal economic activity. Surprisingly little is known about the short run determinants of currency demand.

Recently, modeling and forecasting of daily cash demand in cash supply chains has gained prominence by two papers. Cabrero et al. (2009) consider the daily series of banknotes in circulation in the context of liquidity management of the European Central Bank (ECB). The authors analyze and compare the forecasting accuracy of an ARIMA model with a structural time series model based on the Root Mean Squared Forecast Error (RMSE) and the forecasting accuracy test by Diebold and Mariano (1995). Their empirical results suggest the two econometric models explain large parts of the variations in the daily series with the ARIMA model yielding a lower accuracy over forecasting horizons of up to 4 days and higher accuracy for forecasting horizons of more than 4 days compared to the structural time series model. Brentnall et al. (2008) take a different approach by studying the temporal process of cash withdrawals for individuals. The developed point process model describes the occurrence of individual cash withdrawals over time. Both studies provide strong evidence for seasonality and calendar effects in the series of daily demand for cash, but do not account for the potentially interrelated nature of demand in cash supply chains. The present study differs from previous work by (1) considering co-variability and (2) by focusing on points of cash withdrawal, not individual customers or macroeconomic aggregates.

Supply chain management is an approach to efficiently integrate the upstream and downstream relationships to provide goods in the right quantities, to the right locations, at the right time in order to minimize system wide costs while meeting service level requirements (Simchi-Levi et al., 2007). Hence, cash supply chain management adopts a system wide view by emphasizing the interrelated nature of currency demand on a microeconomic level. Integration is rendered difficult by the variability in the supply chain due to the uncertainty of demand.

Clearly, independent forecasts isolate individual demand series and potentially amplify disturbances. One commonly suggested way to improve forecasting accuracy of demand in supply chains is the use of central information for joint forecasting (Aviv, 2002). However, empirical evidence regarding the benefits of this approach is to be presented. As such, the present study is the first to capture seasonality, calendar effects and cross-variable dynamics in daily cash demand series in cash supply chains.

Following the rationale outlined above, this study investigates forecasting accuracy and the potential of joint forecasting in cash supply chains by contrasting a vector time series model and a seasonal ARIMA model. Calendar effects are modeled using exogenous variables. The impact of co-variability is illustrated for a network of ATMs. Forecasting accuracy is evaluated using daily cash dispense records.

**MATERIALS AND METHODS**

**Data:** The data for this study is obtained from a mid-size European bank with total customer deposits in 2009 of approximately EUR 25 billion and contains cash dispense records of a randomly selected region with 20 ATMs serving approximately 30000 people. Each time series comprises 759 daily cash withdrawals covering the period March 21st 2007 to April 17th 2009. The size of the region provides a reasonable tradeoff between tractability and data requirements. In comparison, the same number of ATMs serves on average 15000 people in Spain, 29000 people in France, 32500 in Germany, 64000 people in Finland or 26000 people in the euro area (European Central Bank, 2009).

The data is analyzed first using a visual inspection of the time series plots. The data displays no long-run trend, but clearly a seasonal pattern and cyclical behavior in daily cash withdrawals. Most noticeable in all sample time series is that daily cash withdrawals seem to increase in December and drop in the months January and February. Daily, weekly and monthly patterns appear to be stable over time. However, each series seems to be affected by local events like festivals, promotions and technical failures of dispense devices, which may explain some of the unusual spikes and drops not related to calendar effects. These events alter the flow of people and consequently impact cash withdrawals. Figure 1 depicts daily cash withdrawals of one ATM in the data set and is exemplary for the other time series considered in this study.

Further evidence on seasonality and calendar effects in daily cash withdrawals is presented by a grouped box plot analysis. The box plot provides a graphical representation of summary statistics. The values are the minimum, 1st, 5th and 25th percentile, median, mean, 75th, 95th and 99th percentile and the maximum. Figure 2 shows day-of-the-week, month-of-the-year and holiday effects. The box plot displays larger daily cash withdrawals during the months March, October and December. Cash withdrawals are lower for the months January, February and April as well as on public holidays.
In addition, the box plot reveals variations in daily cash withdrawals over the course of a week. Daily cash withdrawals tend to reach their peak on Thursdays and respective lows on Mondays with only small fluctuations occurring between Saturday and Tuesday. Furthermore, Figure 3 indicates a day-of-the-month effect with relatively more cash being dispensed towards the end of the month.

The findings seem to confirm earlier reports of seasonality and calendar effects with respect to daily demand for cash. Cabrero *et al.* (2009) identify weekly, monthly and annual seasonal patterns, which resemble a certain regularity in payments and behavior. Similarly, the trading day effect results in an increase in the amount of banknotes in circulation just before the weekend that reverses after the weekend. Likewise, the number of banknotes in circulation decreases before the middle of the month and increases towards the end of the month. The authors further report a strong impact of holidays on the demand for cash.

Brentnall *et al.* (2008) reach similar conclusions for the occurrence of cash withdrawals for individuals and report a weekly cyclical pattern as well as holiday effects. Although the occurrence rate of cash withdrawals seems to differ largely among individuals, the long-term rate is fairly constant. However, no information regarding the amounts per withdrawal is provided.

In order to forecast demand in cash supply chains, a framework is needed that captures seasonal and calendar effects of the individual time series as well as possible co-variability among the time series. Next, two approaches are contrasted and presented: a seasonal ARIMA model and a vector time series model.

**Seasonal ARIMA model:** The seasonal Box-Jenkins model considered in this study is generalizing the ARIMA model to time series containing stochastic seasonal periodic components (Box *et al.*, 2008). The multiplicative seasonal ARIMA model is said to be of the type SARIMA \((p,d,q)\times(P,D,Q)\):

\[
\phi(B)\Phi(B)\nabla^d\nabla_s^q y_t = \Theta(B)\theta(B)a_t,
\]

Where:

\(\phi(B)\) = The regular autoregressive polynomial of order \(p\)

\(\Phi(B)\) = The seasonal autoregressive polynomial of order \(P\)

\(\theta(B)\) = The regular moving average polynomial of order \(q\)

\(\Theta(B)\) = The seasonal moving average polynomial of order \(Q\)

The differentiating operator \(\nabla^d\) and seasonal differentiating operator \(\nabla_s^q\) eliminate non-seasonal and seasonal non-stationarity. The term \(a_t\) follows a white noise process and \(s\) defines the seasonal period.
The joint data generation process may capture additionally variables determined outside the system, such as calendar effects. Respectively, the VAR(p) model with exogenous variables is referred to as VARX(p,q):

\[ Y_t = W_0 + \sum_{i=1}^{p} A_i Y_{t-i} + \sum_{j=1}^{q} B_j X_{t-j} + U_t \] (10)

Where:
- \( B_j = (K \times M) \) coefficient matrices
- \( X_t = M \)-dimensional vector of exogenous variables

The order of the VARX model concerning endogenous and exogenous variables is determined by the parameters \( p \) and \( q \).

Next, the steps involved in specifying, estimating and testing of the models outline above are described. Adequate models are identified using the first 731 observations \( t = \{1,...,731\} \) of each time series. Consequently, the holdout sample covers the last 28 observations of each time series preserving two full years of data for in-sample analysis and model estimation.

**Seasonal ARIMA model specification**: Specification of the SARIMA \((p,d,q) \times (P,D,Q)_s\) model is based on enumerative search following the procedure presented in Wei (2005). The enumerative search in this study considers models of the type \( p \in \{0,1,2\}, q \in \{0,1,2\}, P \in \{0,1,2\}, Q \in \{0,1,2\}, d \in \{0,1,2\}, D \in \{0,1,2\} \) and the seasonal period \( s = 7 \). The order of integration is selected using non-seasonal and seasonal unit root tests (Ghysels and Osborn, 2001). The model choice among the class of adequate models for each time series is based on the 1-step ahead forecasting error using the Mean Absolute Percentage Error (MAPE).

Table 1 depicts the single best performing SARIMA model for each of the 20 ATMs for the in-sample period \( t = \{1,...,731\} \). All series, except ATM19 are seasonally integrated and require seasonal differencing of order one. In fact, the SARIMA \((0,0,0) \times (0,1,1)_7\) model is the single best performing model for 14 out of 20 ATMs. The selected models for the remaining series include additionally non-seasonal differencing of order one and moving average and autoregressive polynomials. The SARIMA \((0,1,1) \times (0,1,1)_7\) model, which resembles the well know airline model, is selected for ATM1 and ATM16.

Extensions of the SARIMA model include an intervention function and concern the impact of calendar effects on the demand for cash. Adequate models contain only parameters that are statistically significant (\( p<0.05 \)).
Selection of the best performing model is again based on the MAPE. Exogenous variables for day of the week and month of the year allow to capture calendar effects that result from the change of the relative position from year to year, which otherwise would not be taken into account. Other calendar effects captured by exogenous variables identify public or bank holidays, the day before a public or bank holidays and end of the month.

**Vector time series model specification:** Specification of the vector time series model involves the selection of the VAR order. Obviously, the true order \( p \) of the observed data generation process is unknown. An intuitive choice of the VAR order \( p \) for the given empirical data with a weekly cyclical behavior is \( p = 7 \). The literature suggests a range of criteria to avoid fitting VAR models with unnecessarily large orders (Lütkepohl, 2005). The Hannan-Quinn Information Criterion (HQIC) and the Schwarz-Bayes Information Criterion (SBIC) provide consistent order selection criteria. Both, the HQIC and the SBIC suggest a VAR model of order \( p = 1 \). The Akaike Information Criterion (AIC) and the Final Prediction Error (FPE) criterion minimize the forecast MSE. However, the latter two suggest a VAR model of order \( p = 2 \). Both, AIC and FPE are known to asymptotically overestimate the true order. Hence, VAR models of order \( p = 1 \) are considered for further evaluation. The whiteness of the error terms is confirmed by a portmanteau test and LM test.

VARX models consider in addition the same calendar effects as specified for the SARIMA model with exogenous variables. Coupling effects, such as between the day before a public or bank holiday and the actual holiday are captured using separate calendar effects. Hence, a VARX(p,q) model with order \( q = 0 \) is considered, which reduces to a VAR(p) model with exogenous variables, sometimes referred to as VARX(p) model.

Subject to exogenous variables, three alternative models are considered to gain further insight to the impact of calendar effects on forecasting accuracy. The first model represents a model with all calendar effects. The second model only considers the day-of-the-week and month-of-the-year effects, while the third model is restricted to the day-of-the-week effect.

**RESULTS**

Forecasting accuracy is measured using the mean absolute percentage error depicted in (11). Despite its criticism, the MAPE remains a popular and recommended error measure (Hyndman and Koehler, 2006). In this study, the applicability of the measure is not compromised by scale incompatibilities, negative or close to zero values. Hence, there is no compelling reason to apply alternative measures that aim at addressing potential shortcomings for these specific cases. It is however acknowledged that MAPE puts a heavier penalty on positive forecasting errors than on negative forecasting errors. The sometime suggested Symmetric Mean Percentage Error (SMAPE) has instead a heavier penalty on negative forecasting errors and therefore provides no considerable alternative:

\[
\text{MAPE} = \left( \frac{1}{T} \sum_{t=1}^{T} \frac{|Y_t - F_t|}{Y_t} \right) \times 100\%
\] (11)

Forecasting accuracy of the ARIMA model and the vector time series model are compared as well as contrasted with two simple models that were included as benchmark. One is the naive model, which is often considered for this type of comparison. Like a random walk model, it uses the most recent observation as a predictor. The other model, a seasonal naive model, uses the last observation of the same season as predictor. The other model, a seasonal naive model, uses the last observation of the same season as predictor.

The accuracy of each model is measured and compared for the in-sample \( t = \{1,...,731\} \) and the holdout sample \( t = \{732,...,759\} \). In order to obtain more robust results, the holdout period of 28 days is split further in two periods of 14 days. In addition, the one-step ahead non-cumulative error is compared with the cumulative error for the holdout samples \( t = \{732,...,745\} \) and \( t = \{746,...,759\} \).

Results of the various models and sample periods are summarized in Table 2. Clearly, the vector time series models perform best for the in-sample period and outperform the SARIMA models.
Table 2: Forecasting errors (MAPE)

<table>
<thead>
<tr>
<th>Model</th>
<th>In-sample non-cumulative (%)</th>
<th>Holdout I non-cumulative (%)</th>
<th>Holdout II non-cumulative (%)</th>
<th>Holdout I cumulative (%)</th>
<th>Holdout II cumulative (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>SARIMA</td>
<td>41.10</td>
<td>19.06</td>
<td>21.81</td>
<td>2.03</td>
<td>1.59</td>
</tr>
<tr>
<td>SARIMA + Day + Month + Other</td>
<td>30.74</td>
<td>19.23</td>
<td>20.72</td>
<td>4.59</td>
<td>3.92</td>
</tr>
<tr>
<td>VAR(1)</td>
<td>37.65</td>
<td>32.97</td>
<td>31.47</td>
<td>5.75</td>
<td>6.46</td>
</tr>
<tr>
<td>VAR(7)</td>
<td>25.73</td>
<td>20.32</td>
<td>29.77</td>
<td>4.05</td>
<td>5.12</td>
</tr>
<tr>
<td>VAR(1) + Day + Month + Other</td>
<td>26.10</td>
<td>21.65</td>
<td>23.88</td>
<td>7.60</td>
<td>8.63</td>
</tr>
<tr>
<td>VAR(1) + Day + Month</td>
<td>26.77</td>
<td>20.33</td>
<td>24.80</td>
<td>8.39</td>
<td>7.33</td>
</tr>
<tr>
<td>VAR(1) + Day</td>
<td>27.33</td>
<td>20.77</td>
<td>24.94</td>
<td>9.42</td>
<td>5.72</td>
</tr>
<tr>
<td>VAR(7) + Day + Month + Other</td>
<td>23.26</td>
<td>21.57</td>
<td>27.62</td>
<td>4.89</td>
<td>6.55</td>
</tr>
<tr>
<td>VAR(7) + Day + Month</td>
<td>23.90</td>
<td>20.49</td>
<td>28.80</td>
<td>4.31</td>
<td>5.79</td>
</tr>
<tr>
<td>VAR(7) + Day</td>
<td>24.23</td>
<td>20.23</td>
<td>29.10</td>
<td>4.11</td>
<td>4.82</td>
</tr>
<tr>
<td>Naive</td>
<td>50.12</td>
<td>50.94</td>
<td>51.35</td>
<td>1.66</td>
<td>1.83</td>
</tr>
<tr>
<td>Seasonal naive</td>
<td>34.88</td>
<td>18.22</td>
<td>32.51</td>
<td>2.73</td>
<td>5.54</td>
</tr>
</tbody>
</table>

Notes: Results depict the mean absolute percentage error for the network of 20 ATMs and compare the 12 models for time periods $t = \{1,...,731\}$ (March 21st, 2007-March 20th, 2009), $t = \{732,...,745\}$ (March 21st, 2009-April 3rd, 2009) and $t = \{746,...,759\}$ (April 4th, 2009-April 17th, 2009).

Exogenous variables that capture calendar effects further reduce the errors of SARIMA models and vector time series models. Limiting the calendar effects under consideration to the day-of-the-week and month-of-the-year or simply the day-of-the-week results in larger MAPEs. Nevertheless, these models outperform their counterparts without exogenous variables. The VAR(7) model yields a higher accuracy during the in-sample period than the VAR(1) model. This effect diminishes and reverses for the models including exogenous variables during the holdout period.

However, the vector time series model VAR(7) without exogenous variables results in lower MAPEs than VAR(1) suggesting that an over-specified model can reduce forecasting errors.

The result for the non-cumulative holdout sample is not as unambiguous as for the in-sample period. MAPEs for the holdout period are generally lower, but increase during the second half. This indicates that holiday effects present in $t = \{746,...,759\}$ are only partially accounted for and may in fact be determined by additional factors such as weather and festivals. In fact, the holiday effect is reversed during the holdout period with more cash being dispensed than on an average day or compared to the in-sample period.

The seasonal naive model outperforms the non-seasonal naive model during the in-sample as well as the holdout period. MAPEs of the seasonal naive model are in line with the SARIMA model and vector time series models suggesting that a large part of the variability in the demand series is attributed to seasonality in form of the day-of-the-week effect.

Moreover, cumulative forecasts result in exceptionally low MAPEs suggesting that most forecasting errors are offset over a period of 14 days. Cumulative 14 days forecasts using the naive method match the performance of the SARIMA model without exogenous variables despite the non-cumulative errors being more than twice as large. Joint forecasting and exogenous variables do not further reduce MAPEs for cumulative forecasts.

**DISCUSSION**

In this study, a framework for investigating the overall forecasting accuracy of cash demand and the potential of joint forecasting in cash supply chains is presented. The approach is motivated by advances in supply chain management. The applied methodology considers the potentially interrelated nature of individual daily cash demand series combined with seasonality and calendar effects. In particular, a vector time series model and a seasonal ARIMA model were compared using naive methods as benchmark.

The vector time series model captures the dynamics and co-variability that characterizes a multi-stage supply chain. From a theoretical perspective, such a model is expected to yield more accurate forecasts, thereby enabling integration of upstream and downstream relationships and improving the efficiency within the cash supply chain. The approach is illustrated for a network of ATMs.

Forecasting accuracy was measured and compared using MAPE. Both, the SARIMA model and the vector time series model resulted in lower MAPEs than the respective benchmark models. Results showed that overall forecasting accuracy is high with a mean absolute percentage error of approximately 20 percent. Particularly, cumulative forecasts for periods of 14 days are very robust with mean percentage errors of approximately two percent. Specification of models revealed strong and consistent seasonal patterns that led to the SARIMA $(0,0,0)\times(0,1,1)_7$ model being the single best performing SARIMA $(p,d,q)\times(p,D,Q)_s$ model. The vector time series model resulted in smaller MAPEs than the SARIMA model during the in-sample period. However, the joint forecasting approach did not yield a higher accuracy.
than the independent forecasting approach for the holdout sample. Exogenous variables capturing calendar effects accounted for part of the variability in daily cash demand series. Results suggest further that other factors such as weather and local events affect the demand for cash and may even dominate calendar effects.

Practical implications of this study concern the ability to forecast currency needs in order to manage cash supply chains more efficiently. More accurate forecasts enable cost savings by reducing excess cash holdings as well as by cutting the number of emergency replenishments needed to prevent cash outs. For example, a EUR 10000 reduction of the average stock held by a bank branch or ATM results in annual cost savings of EUR 400, given the cost of capital is four percent per annum. The potential economic impact is large considering the entire euro area with 190886 commercial bank branches and 249705 ATMs (European Central Bank, 2009) or the MasterCard network with more than 1 million ATMs worldwide (MasterCard, 2009).

CONCLUSION

Results confirm the potential of advanced forecasting techniques for cash supply chains and help explain variability in daily cash demand. Reduced forecasting errors result in lower demand uncertainty and enable supply chain partners to coordinate replenishments across the cash supply chain. However, the potential for information sharing seems to be limited to this role and does not necessarily translate into more accurate forecasts based on joint forecasting.

Limitations of this study arise from the simple two-stage supply chain framework and the relatively small amount of data available. Long-run time series contain more observations of calendar effects such as holidays and may provide further insights. In addition, cumulative forecasts are very robust, since forecasting errors are offset over a period of 14 days. Similar effects may explain, why the joint forecasting model does not yield substantially better forecasts. In fact, anomalies in one series may be passed on to all other series increasing forecasting errors system-wide. However, limitations linked to the applied error measure, the mean absolute percentage error, can be regarded minimal given the nature of the data.

Results of this study provide preliminary evidence and call for further investigation of joint forecasting in a multi-stage supply chain spanning more than two stages. Further research will have to address additional exogenous variables such as weather and festivals that may dominate certain calendar effects. Advanced forecasting models may additionally provide the link to scenarios and risk models in cash supply chain management.

REFERENCES


