ADP FRANCE NATIONAL EMPLOYMENT REPORT

METHODOLOGY

A more human resource.®

Introduction and Motivation

This document describes the ADP France National Employment model for monthly forecasts of France non-farm employment. While the official non-farm payroll employment data is published by the French National Institute of Statistics and Economic Studies (INSEE) on a quarterly basis, the estimates of employment changes for France are generated and reported on monthly basis based on ADP's transactional payroll data in France and prior to the official monthly unemployment rate number for the country.

This redevelopment of the model is introduced due to several major factors. The most notable ones include changes in INSEE's methodology for the estimation of payroll employment, for both flash and quarterly estimates; attempt to cover larger extent of employment numbers permitted by ADP coverage of all major employment sectors and finally to incorporate a more suitable benchmarking and data cleaning techniques.

In the model set-up, the ADP data contains all the 16 main industry groups reported by INSEE each quarter. This is driven *inter alia* by significant methodological changes introduced by INSEE in 2017. The data used for modeling is a panel data containing an average of about 70,000 firms that vary in size and industry. For each firm the number of active employees is reported for a specific day of the month which is defined as the date of the most recent payroll process. The time period ranges from 2009 onwards, however, the modeling sample starts in 2013 to allow for better statistical inference.

Model Set-up

In the current model and procedure set-up, several things changed. First, the data cleaning process was amended so that a matched pairs algorithm is respected always. Also, following a close inspection of the ADP payroll data, critical values and limits used for month-on-month employment changes in raw data were adjusted to account for the presence of outliers. Further, given the increasing share of the number of firms from the agricultural sector reporting payrolls, the sector was also included among the modeled sectors, mainly due to the potential of spillover effects into other sectors. Likewise, interpolation techniques for quarterly INSEE numbers were enhanced and the end-of-month biases reduced. As indicated earlier, the whole available history of official INSEE numbers was revised with the new reporting standard.

Following a revised data cleaning technique, the sample was restricted to 2013 onwards (until the last available observation). Several new drivers were tested against sectoral employment and added to the model. After adjusting and redefining individual VARX variable constraints, the model was split into two parts: The first part has the form of a standard vector autoregression for selected variables, like the set-up of the original model. The second part of the model introduces individual OLS equations inclusive of autoregressive terms for specific industries. The sectors include trade, business services, agriculture, public services, household services, and manufacturing. However, the two models should be viewed as challenger models, while statistical and economic inference is done with every update about which model set-up produces more satisfactory results and forecasts.

Data Description and Analysis

In this section data cleaning technique is analyzed. As mentioned previously, ADP data take form of a monthly frequency panel data containing about 70,000 firms of different sizes and belonging to different industries. The codes and names of the industries selected for the model are presented in the tables below.

Table 1: Codes and names of the industries selected for modeling

Code	Description	Aggregated categories
AZ	Agriculture, Forestry, and Fishing	Agriculture, Forestry, and Fishing
DE	Energy, Water and Waste	Energy, Water and Waste
C1	Food and Drinks	
C2	Coke and Refined Petroleum	
C3	Machinery and Equipment Goods	Manufacturing
C4	Transport Equipment	
C5	Other Industrial Goods	
FZ	Construction	Construction
GZ	Trade	Trade
HR	Transportation	Transportation
IZ	Accommodation and Food Services	Accommodation and Food Services
JZ	Information and communication	Information and Communication
KZ	Financial Services	Financial Services
LZ	Real Estate Services	Real Estate Services
MN	Business Services	Business Services
OQ	Public Services	Public Services
RU	Household Services	Household Services

For each firm the number of active employees is reported at a specific day of the month which is defined as the date of the most recent payroll process and report. The time period used for modeling ranges from 2013 onwards. Since the data panel is unbalanced, i.e. it does not contain the same amount of observations at each point in time and/or not all characteristics are observable in all data points in the time series, interpolation is used to handle the missing observations. Such approach is used to keep the matched pairs algorithm in place and to fill in one-month gaps or two-month gaps in data or compensate for changes in monthly sample sizes.

Once the missing observations are taken care of, the outliers are removed to smooth out monthly percentage changes. After inspecting the monthly variation in the data, the criterion to remove the outliers is customized for different firm sizes to capture the reality and minimize the measurement error if any and differ also depending on whether we observe increase or a decrease.

VARX Model

We employ vector autoregressive methodology to model and forecast monthly changes in employment. Vector autoregressive models (VARX) are widely used for analyzing multivariate time series. Typical VARX models treat all variables as endogenous following Sim's (1980) critique on the ad hoc exogeneity assumption of macroeconomic models. VARX models can incorporate restrictions, including the exogeneity of some of the variables. They can also be amended to include deterministic terms and exogenous variables. VARX models are used for modeling and forecasting the dynamic behavior of economic and financial time series. They can also be used for structural inference and policy analysis. The basic VARX(p) model is defined by the following equation

$$y_t = A_1 y_{t-1} + \dots + A_p y_{t-p} + u_t ,$$

where A_i are (K x K) coefficient matrices and u_t is a K dimensional white noise process with $E(u_t) = 0$ and $E(u_t u_t^T) = \sum u$. The VARX(p) model is a stationary process with time invariant mean, variance and covariance structure. Besides the lags of endogenous variables VARX model can also include exogenous regressors on the right-hand side and in this case VARX is represented by the following equation

$$y_t = A_1 y_{t-1} + \dots + A_p y_{t-p} + B_1 X_{t-1} + \dots + B_p X_{t-p} + u_t,$$

where X_t represents the vector of exogenous variables.

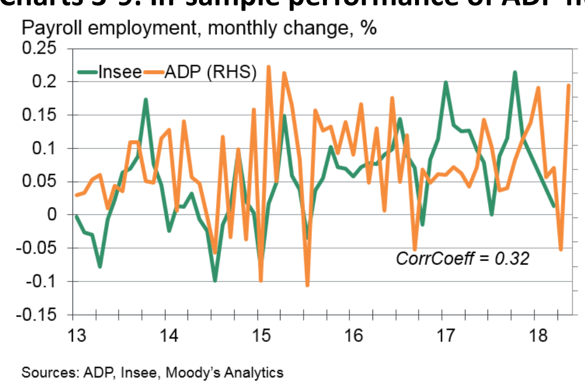
The coefficients of the VARX(p) process can be estimated by applying OLS to each of the equations. The OLS estimator is identical to the GLS estimator if there are not restrictions in the parameters. For a normally distributed Gaussian Process $y_t \sim N(0, \sum u)$, the OLS estimator is also identical to the ML estimator. The usual statistical inference procedures can be applied if the process is stable. If there are integrated variables so that $y_t \sim I\,(1)$, then the process is unstable and the variables may be cointegrated. The OLS / ML method can still be applied for estimating the model parameters but the usual t-statistics and F-statistics can lead to misleading conclusions when used for hypothesis testing. VARX can also be represented in structural form if restrictions are imposed on model parameters, that is, elements of A matrix, B matrix or both matrices in the equation above.

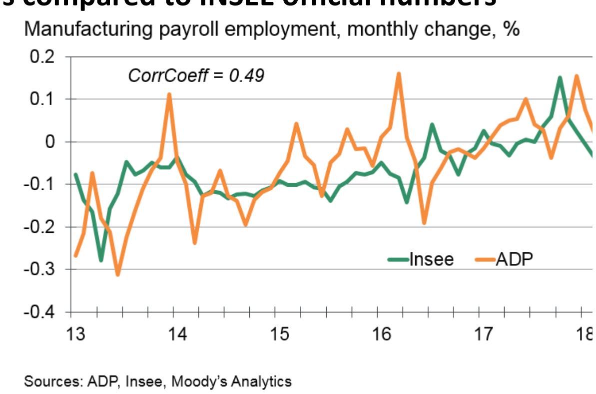
In our setting using VARX model allows to incorporate spillover and feedback effects on employment among different industries and to measure the common movement of employment in different industries with backward-looking and forward-looking business cycle indicators. We start with estimating unrestricted VARX with one lag in which endogenous variables are represented by month-on-month changes in 11 industries. We choose to aggregate employment in C1, C2, C3, C4 and C5 sectors into one sector which we call total manufacturing. The remaining industries included in our model correspond to INSEE industries.

Results

On the charts below, we plot fitted values of month on month percentage changes in ADP employment for selected industries together with percentage changes in INSEE employment. As can be seen, the results are in most cases satisfactory and represent an improvement compared to the previous set-up. The representations in the charts below refer to in-sample behavior.

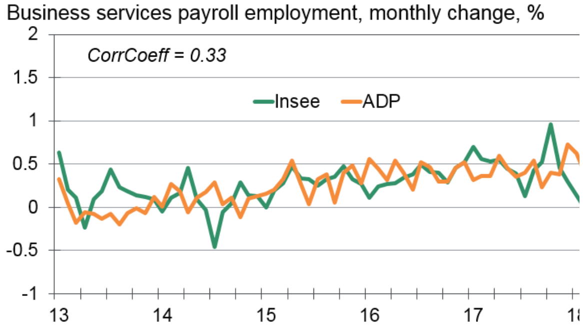
Charts 3-9: In-sample performance of ADP numbers compared to INSEE official numbers





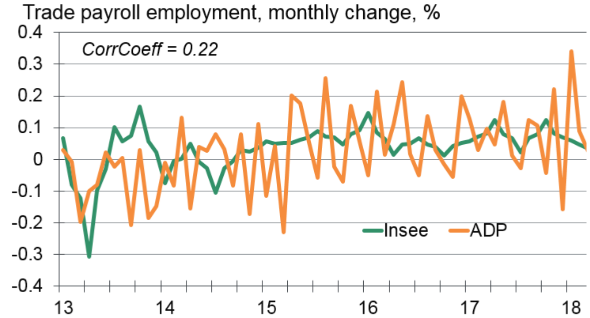
A more human resource.®

Business services payroll employment, monthly change, %



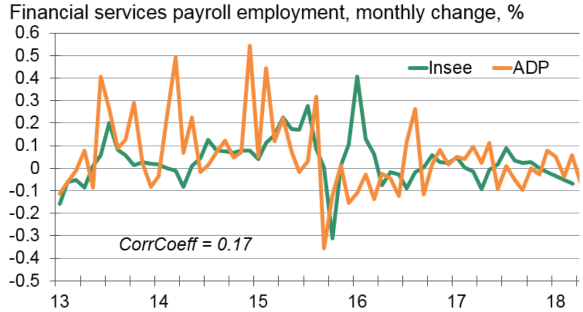
Sources: ADP, Insee, Moody's Analytics

Trade payroll employment, monthly change, %



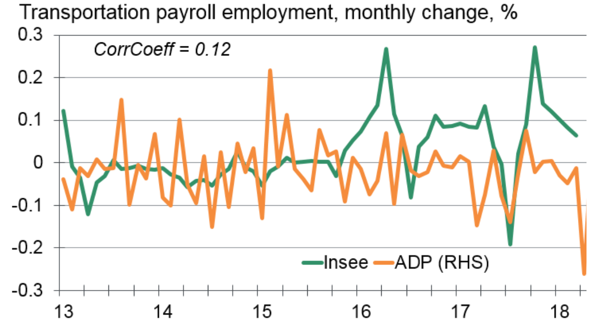
Sources: ADP, Insee, Moody's Analytics

Financial services payroll employment, monthly change, %



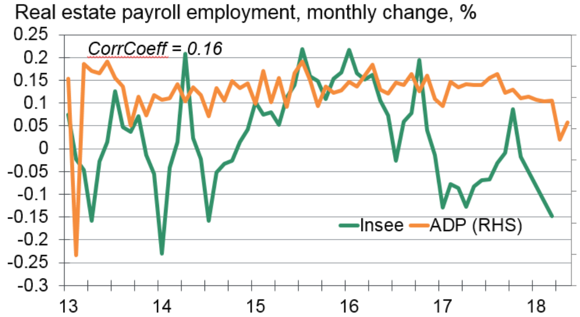
Sources: ADP, Insee, Moody's Analytics

Transportation payroll employment, monthly change, %



Sources: ADP, Insee, Moody's Analytics

Real estate payroll employment, monthly change, %



Sources: ADP, Insee, Moody's Analytics