

# SVAR-based supply and demand decomposition of inflation

Pablo Barrio, Benjamin Pipaud, Pierre Rouillard

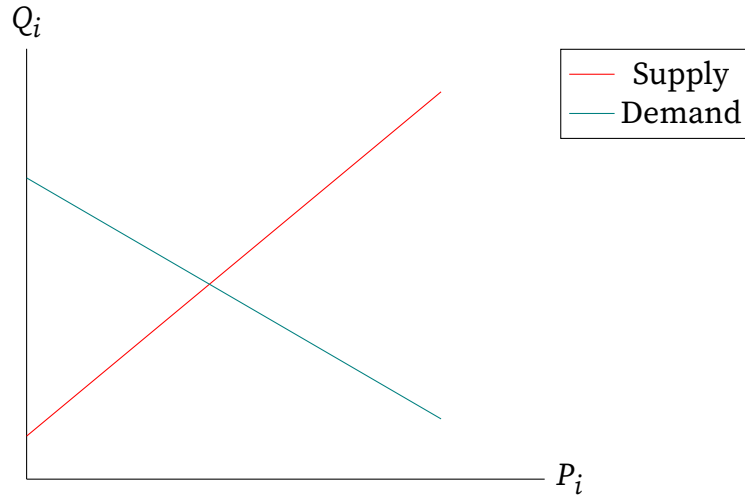
<DEPLOYED APP>

1. Methodology and Data	1
1.1. Basis of <i>Shapiro (2022)</i> . . . . .	1
1.2. Data for European countries . . . . .	3
2. Estimation	4
2.1. Classification from <i>Shapiro (2022)</i> . . . . .	4
2.1.1 Baseline classification . . . . .	4
2.1.2 Robustness check : alternative classifications . . . . .	6
2.2. Classification from <i>Sheremirov (2022)</i> . . . . .	6
2.3. Proposed classification rule . . . . .	7
3. Results	10
3.1. Results for France . . . . .	10

# 1. Methodology and Data

## 1.1. Basis of Shapiro (2022)

The methodology proposed by Adam Hale Shapiro in his 2022 paper "Decomposing Supply and Demand Driven Inflation" takes its foundations in the work of Jump and Kohler (2022). The core assumption of the framework is that for each sector/section ( $i$ ) of the inflation decomposition, an upward sloping in price supply curve and a downward demand curve can be assumed.



In theory, assuming  $\sigma_i > 0$  and  $\delta_i > 0$ , we have:

$$\text{Supply : } Q_i = \sigma_i \cdot P_i + \alpha_i$$

$$\text{Demand : } P_i = -\delta_i \cdot Q_i + \beta_i$$

Shocks are then defined as vertical movements of the curves :

$$\text{Supply shock : } \varepsilon_t^s = \Delta \alpha_i = (Q_{i,t} - \sigma_i P_{i,t}) - (Q_{i,t-1} - \sigma_i P_{i,t-1})$$

$$\rightarrow \varepsilon_t^s = \Delta Q_{i,t} - \sigma_i \Delta P_{i,t}$$

$$\text{Demand shock : } \varepsilon_t^d = \Delta \beta_i = (P_{i,t} + \delta_i Q_{i,t}) - (P_{i,t-1} + \delta_i Q_{i,t-1})$$

$$\rightarrow \varepsilon_t^d = \Delta P_{i,t} + \delta_i \Delta Q_{i,t}$$

We can easily show that the two previous equations lead to :

$$\begin{cases} \Delta Q_{i,t} &= \frac{1}{1+\delta_i}(\varepsilon_t^s + \sigma_i \cdot \varepsilon_t^d) \\ \Delta P_{i,t} &= \frac{1}{\sigma_i(1+\delta_i)}(\sigma_i \cdot \varepsilon_t^d - \delta_i \cdot \varepsilon_t^s) \end{cases}$$

As we assume  $\sigma_i > 0$  and  $\delta_i > 0$ , we can derive expected comovements between the two variables following supply and demand shocks.

$$\begin{cases} \text{Supply shock } \Delta^+ \varepsilon^s & : \Delta^+ Q \quad \& \quad \Delta^- P > \text{Negative comovements (S)} \\ \text{Demand shock } \Delta^+ \varepsilon^d & : \Delta^+ Q \quad \& \quad \Delta^+ P > \text{Positive comovements (D)} \end{cases}$$

Consider the following structural VAR of (arbitrary) order  $p$  (dropping  $i$  indices). Let

$$z_t = \begin{bmatrix} \Delta Q_t \\ \Delta P_t \end{bmatrix} \text{ and the structural shocks } \varepsilon_t = \begin{bmatrix} \varepsilon_t^s \\ \varepsilon_t^d \end{bmatrix} :$$

$$A \cdot \begin{bmatrix} \Delta Q_t \\ \Delta P_t \end{bmatrix} = \mu + \sum_{i=1}^p A_i \cdot z_{t-i} + \begin{bmatrix} \varepsilon_t^s \\ \varepsilon_t^d \end{bmatrix}$$

With  $v_t$  the residuals of the estimated reduced-form VAR( $p$ ) we should have  $v_t = A^{-1} \cdot \varepsilon_t$

Let  $A$  satisfy  $A \equiv \begin{pmatrix} 1 & -\alpha \\ \beta & 1 \end{pmatrix}$  with  $\alpha, \beta > 0$ , it follows that  $A^{-1} = \frac{1}{1+\alpha\beta} \begin{pmatrix} 1 & \alpha \\ -\beta & 1 \end{pmatrix}$

Omitting the  $t$  indices we have :

$$\begin{aligned} \begin{bmatrix} v^s \\ v^d \end{bmatrix} &= A^{-1} \cdot \begin{bmatrix} \varepsilon^s \\ \varepsilon^d \end{bmatrix} \\ \begin{bmatrix} v^s \\ v^d \end{bmatrix} &= \frac{1}{1+\alpha\beta} \begin{pmatrix} 1 & \alpha \\ -\beta & 1 \end{pmatrix} \begin{bmatrix} \varepsilon^s \\ \varepsilon^d \end{bmatrix} \end{aligned}$$

As  $\frac{1}{1+\alpha\beta} > 0$ , we finally have:

$$\begin{cases} v^s & \propto \varepsilon^s + \alpha \cdot \varepsilon^d \\ v^d & \propto -\beta \cdot \varepsilon^s + \varepsilon^d \end{cases}$$

Which leads to:

$$\begin{cases} \varepsilon^s > 0, \varepsilon^d > 0 & \Rightarrow v^s > 0 \\ \varepsilon^s < 0, \varepsilon^d < 0 & \Rightarrow v^s < 0 \\ \varepsilon^s < 0, \varepsilon^d > 0 & \Rightarrow v^d > 0 \\ \varepsilon^s > 0, \varepsilon^d < 0 & \Rightarrow v^d < 0 \end{cases}$$

Looking at the signs of  $v^s$  and  $v^d$  we can then back-track and derive the signs of  $\varepsilon^s$  and  $\varepsilon^d$ . For example when  $v^s > 0$  and  $v^d < 0$  we have  $\varepsilon^s > 0$ . We derive:

$$\begin{cases} v^s > 0, v^d < 0 \Rightarrow \varepsilon^s > 0 & + \text{Supply shock} & (1) \\ v^s < 0, v^d > 0 \Rightarrow \varepsilon^s < 0 & - \text{Supply shock} & (2) \\ v^s > 0, v^d > 0 \Rightarrow \varepsilon^d > 0 & + \text{Demand shock} & (3) \\ v^s < 0, v^d < 0 \Rightarrow \varepsilon^d < 0 & - \text{Demand shock} & (4) \end{cases}$$

Thus, from (1) and (2) we see that reduced-form errors of opposite sign (negative comovements of  $P$  and  $Q$ ) are associated with demand shocks, which is consistent with theory (S). Similarly, from (3) and (4) reduced-form errors of the same sign (positive comovements of  $P$  and  $Q$ ) are associated with demand shocks, again consistent with theory (D).

Hence, assuming that  $A$  in the structural VAR is of the form  $\begin{pmatrix} 1 & -\alpha \\ \beta & 1 \end{pmatrix}$  with  $\alpha, \beta > 0$ , or

more generally  $A = \begin{pmatrix} a_{11} > 0 & a_{12} < 0 \\ a_{21} > 0 & a_{22} > 0 \end{pmatrix}$ , thus **ensures** that expected structural shocks' effects on the covariates derived from theory are consistent in the model. We can then infer aforementioned effects from the reduced-form residuals that are in practice derived from the estimation of a VAR( $p$ ) model.

## 1.2. Data for European countries

In the original paper from Shapiro the data used are price, quantity and expenditures from the personal consumption expenditure (PCE) data with fourth level of disaggregation of the Bureau of Economic Analysis (BEA). However it is more complicated when considering European countries. Indeed, similar quantity series are not readily available

for each of the 4-digit COICOP classification component of the HICP.

To tackle this issue, we follow the indication provided by Eduardo Goncalves and Gerrit Koester in the ECB Economic Bulletin 7 of 2022. We consider the turnover series from the short-term statistics data of the ECB as proxy for demand, and once deflated as proxy for the quantity series. The main issue that remains is that the classification of sectors in the HICP (COICOP) is different from the classification of sectors used in the short-term statistics data (NACE Rev.2). We then try to match as best as possible each COICOP sector to a NACE Rev.2 counterpart. Out of the 94 4-digit COICOP sectors, we managed to create 75 pairs. On average that is around 80% of the overall HICP that has been matched. In what follows we refer for each sector the 'quantity index' as the real matched turnover series (CPI deflated). The sectors that have not been matched will thus make up the *unclassified* part of total inflation in the proposed charts.

## 2. Estimation

### 2.1. Classification from *Shapiro (2022)*

#### 2.1.1. Baseline classification

With our price-quantity pairs matched for 75 sectors, we run the following monthly VAR for each of the sectors denoted by index ( $i$ ) :

$$\begin{bmatrix} \Delta q_{i,t} \\ \Delta p_{i,t} \end{bmatrix} = \sum_{j=1}^p A_j \cdot \begin{bmatrix} \Delta q_{i,t-j} \\ \Delta p_{i,t-j} \end{bmatrix} + \begin{bmatrix} v_{i,t}^q \\ v_{i,t}^p \end{bmatrix}$$

Where  $\Delta q_{i,t}$  and  $\Delta p_{i,t}$  are the demeaned first-difference of the log-transformed quantity and price indices ( $x_{i,t} = \log(X_{i,t})$ ). This assumes  $\Delta q_{i,t}$  and  $\Delta p_{i,t}$  are second-order stationary, which seems to be the case.

In our implementation of the methodology, we let the order  $p$  of the VAR be automatically selected via minimisation of a statistical criterion (AIC or BIC) for each sector instead of imposing  $p = 12$  as done by Shapiro. Even though he explains the classification seems to be robust to the lag order selection we would rather let it be sector-specific than adding another constraint on the methodology.

Note that we said 75 sectors had matched proxy quantity series, but we cannot run 75 VAR. Indeed, even though we managed to match the series we still had some data

missing in the series themselves. Thus, for each zone in our study (France, Germany, Spain and EU27) we flag sectors that have data problems for either the price or quantity series. A sector being flagged mainly stems from the latest data available being earlier than 2023 (some stop in 2022 or 2020 and some even in 2015). Overall we run regressions on 52 (EU27) to 72 (France) sectors.

The estimated reduced-form residuals  $(v_t^q, v_t^p)$  (equivalent to  $(v_t^s, v_t^d)$  in the previous section) are then used to label each month ( $t$ ) for each sector ( $i$ ) using the following rule

$$\begin{aligned} \mathbb{1}_{i, su p(+)}(t) &= \begin{cases} 1 & \text{if } v_{i,t}^q > 0, v_{i,t}^p < 0 \\ 0 & \text{otherwise} \end{cases} \\ \mathbb{1}_{i, su p(-)}(t) &= \begin{cases} 1 & \text{if } v_{i,t}^q < 0, v_{i,t}^p > 0 \\ 0 & \text{otherwise} \end{cases} \\ \mathbb{1}_{i, dem(+)}(t) &= \begin{cases} 1 & \text{if } v_{i,t}^q > 0, v_{i,t}^p > 0 \\ 0 & \text{otherwise} \end{cases} \\ \mathbb{1}_{i, dem(-)}(t) &= \begin{cases} 1 & \text{if } v_{i,t}^q < 0, v_{i,t}^p < 0 \\ 0 & \text{otherwise} \end{cases} \\ (1) \quad \Rightarrow \mathbb{1}_{i, su p}(t) &= \mathbb{1}_{i, su p(+)}(t) + \mathbb{1}_{i, su p(-)}(t) \\ (2) \quad \Rightarrow \mathbb{1}_{i, dem}(t) &= \mathbb{1}_{i, dem(+)}(t) + \mathbb{1}_{i, dem(-)}(t) \end{aligned}$$

*From now on, at date  $t$  experiencing a demand shock or being demand-driven or being labeled demand-driven all mean that  $\mathbb{1}_{i, dem}(t) = 1$  (vice versa with supply).*

It follows that each month, overall month-over-month inflation  $\pi_{t,t-1}$  can be devided into supply and demand driven components. As the overall CPI is a weighted average of its sectorial inflation  $\pi_{i,t,t-1}$  we denote  $\omega_{i,y}$  the weight of sector ( $i$ ) at year  $y$  in the CPI. In the following equation  $y_t$  means the year of month  $t$ .

$$\begin{aligned} \pi_{t,t-1} &= \sum_i \mathbb{1}_{i, su p}(t) \omega_{i, y_t} \pi_{i,t,t-1} + \sum_i \mathbb{1}_{i, dem}(t) \omega_{i, y_t} \pi_{i,t,t-1} \\ \pi_{t,t-1} &= \pi_{t,t-1}^{su p} + \pi_{t,t-1}^{dem} \end{aligned}$$

We present results for year-over-year inflation, which is simply given by:

$$\begin{cases} \pi_{t,t-12}^{sup} &= \sum_{k=0}^{11} \pi_{t-k,t-k-1}^{sup} \\ \pi_{t,t-12}^{dem} &= \sum_{k=0}^{11} \pi_{t-k,t-k-1}^{dem} \end{cases}$$

### 2.1.2. Robustness check : alternative classifications

Shapiro introduced several alternative classification methods which we replicated to assess whether the methodology was robust enough. In our version, as the lag order for each VAR is automatically selected (and if we believe the selected lag is correct) there should not be specification error.

We implement the the two alternative methods that are the smoothed-errors classification (to offset possible measurement errors) and the parametric approach. In the latter, each month a sector is no longer deterministically classified as experiencing a supply or a demand shock but there rather is a probability  $p_t^K$  that it is experiencing a shock of type  $K \in \{\text{supply}, \text{demand}\}$ .

The smoothed-errors classification amounts to considering  $\sum_{k=0}^j v_{i,k}^q$  and  $\sum_{k=0}^j v_{i,k}^p$  instead of simply  $v_{i,t}^q$  and  $v_{i,t}^p$  in the aforementioned classification. The parameter  $j$  is an hyperparamter, set to be the same for each sector. In his paper Shapiro tries different values (ranging from 1 to 3) meaning for each month  $t$  it is the sign of the (average)  $j$  previous months residuals that matter.

### 2.2. Classification from Sheremirov (2022)

The algorithm proposed by Sheremirov is as follows, for each sector  $i$  and month  $t$  :

1. Compute year-over-year inflation and (real) consumption/demand growth :  $\pi_{i,t}, c_{i,t}$
2. Classify the sector  $i$  as experiencing a demand shock at month  $t$  if both inflation and consumption growth are over their 2001-2019 respective average  $\tilde{\pi}_i, \tilde{c}_i$  :

$$\mathbb{1}_{i,dem}(t) = \begin{cases} 1 & \text{if } (\pi_{i,t} - \tilde{\pi}_i) \cdot (c_{i,t} - \tilde{c}_i) > 0 \\ 0 & \text{otherwise} \end{cases}$$

$$\mathbb{1}_{i,sup}(t) = 1 - \mathbb{1}_{i,dem}(t)$$

3. Classify the shock at month  $t$  as:

$$\begin{aligned} \text{Persistent demand } \mathbb{1}_{i,dem}^{pers}(t) &= \begin{cases} 1 & \text{if } \sum_{k=0}^{11} \mathbb{1}_{i,dem}(t-k) \geq 11 \\ 0 & \text{otherwise} \end{cases} \\ \text{Persistent supply } \mathbb{1}_{i,sup}^{pers}(t) &= \begin{cases} 1 & \text{if } \sum_{k=0}^{11} \mathbb{1}_{i,sup}(t-k) \geq 11 \\ 0 & \text{otherwise} \end{cases} \end{aligned}$$

4. If neither condition is verified, month  $t$  is classified as transitory demand (if  $\mathbb{1}_{i,dem}(t) = 1$ ) or transitory supply.

The Sheremirov methodology was implemented to assess to what extent classifications derived from Shapiro's methods would be consistent with alternative methods, especially one not based on a VAR estimation. However the distinction it provides between transitory and persistent components makes it really interesting for policy-makers even though the classification is solely backward-looking.

### 2.3. Proposed classification rule

On this basis we suggest tweaking the classification outlined by Sheremirov, challenging the set number of months (11) for the persistent classification and the pure backward-looking perspective.

The latter can be a problem as the classification does not rule out persistent demand (or supply) vanishing. Consider a sector  $i$  and assume month  $t - 1$  is not labeled as persistent demand because months  $t - 10$  and  $t - 12$  are not labeled as demand-driven and let month  $t$  be demand-driven. According to Sheremirov's rule, month  $t$  is labeled as persistent demand as we have 11 months classified as demand-driven (only month  $t - 10$  is not). However, if month  $t + 1$  is not demand-driven, persistent demand vanishes after only one month because at least 2 months are not demand-driven conflicting the criterion. Which in our opinion is kind of an awkward situation as the whole purpose of the persitent component is to be persistent and not to instantaneously disappear. Our methodology solves this, as month  $t$  would have never been labeled as persistent demand in the first place due to the forward-looking setting we incorporate to the rule.



On another note, as the first step of the classification below relies on the simple dummy demand-supply classification, we can use both results from the labeling of Shapiro or (step.2 of) Sheremirov. Our aim is to leverage the initial supply-demand classification, decomposing each as transitory-persistent (as Sheremirov did) with both backward and forward-looking perspectives. We add sort of a feedback rule -an ambiguous category- such that the dates for which the forward-looking perspective is not possible (due to unobserved data) are not miss-classified. A month that had been categorized as ambiguous can thus be re-classified once "enough" data have been added.

We propose the following algorithm:

1. Classify month  $t$  for sector  $i$  as experiencing a demand or supply shock: either through Shapiro's rule (1,2) or step.2 of Sheremirov's algorithm (2.2).
2. We introduce the following modified rule, where  $K$  can be debated but we set the baseline as  $K = 3$ :

$$\begin{aligned}
 \text{Persistent demand } \mathbb{1}_{i,dem}^{pers}(t) &= \begin{cases} 1 & \text{if } \mathbb{1}_{i,dem}(t) + \sum_{k=1}^K \{\mathbb{1}_{i,dem}(t-k) + \mathbb{1}_{i,dem}(t+k)\} \geq 2K \\ 0 & \text{otherwise} \end{cases} \\
 \text{Persistent supply } \mathbb{1}_{i,sup}^{pers}(t) &= \begin{cases} 1 & \text{if } \mathbb{1}_{i,sup}(t) + \sum_{k=1}^K \{\mathbb{1}_{i,sup}(t-k) + \mathbb{1}_{i,sup}(t+k)\} \geq 2K \\ 0 & \text{otherwise} \end{cases} \\
 \text{Ambiguous demand } \mathbb{1}_{i,dem}^{abg}(t) &= \begin{cases} 1 & \text{if } t \in [T-K+1, T] \text{ and :} \\ & \mathbb{1}_{i,dem}(t) + \sum_{k=1}^K \mathbb{1}_{i,dem}(t-k) + \mathbb{1}_{t \neq T} \sum_{k=1}^{T-t} \mathbb{1}_{i,dem}(t+k) \geq K + (T-t) \\ & \text{i.e almost satisfies pers. but some date ' } t+k \text{ ' is unobserved!} \\ 0 & \text{otherwise} \end{cases} \\
 \text{Ambiguous supply } \mathbb{1}_{i,sup}^{abg}(t) &= \begin{cases} 1 & \text{if } t \in [T-K+1, T] \text{ and :} \\ & \mathbb{1}_{i,sup}(t) + \sum_{k=1}^K \mathbb{1}_{i,sup}(t-k) + \mathbb{1}_{t \neq T} \sum_{k=1}^{T-t} \mathbb{1}_{i,sup}(t+k) \geq K + (T-t) \\ & \text{i.e almost satisfies pers. but some date ' } t+k \text{ ' is unobserved!} \\ 0 & \text{otherwise} \end{cases}
 \end{aligned}$$

3. If neither condition is verified, month  $t$  is simply classified as transitory demand (if  $\mathbb{1}_{i,dem}(t) = 1$ ) or transitory supply.

The persistent demand(supply) labeling simply asserts that for a given month  $t$ , in the window of the three previous and next months, at least  $2K$  have been classified as experiencing a demand(supply) shock. For instance when  $K = 3$  this checks that for Septembre 2019, between June and Decembre 2019 (7-month window), at least 6 have been labeled as demand-driven (for persistent demand and vice versa).

The *ambiguous* (ambiguous in the sense of undetermined between transitory and persistent) labeling is pretty similar to the persistent but is only applied to the last  $K$  observations. It aims at labeling the months for which we cannot apply the full rule because the forward-looking part simply sums over unobserved data. For instance, when  $K = 3$ , only the last three observations ( $t \in [T - 2, T] = [T - K + 1, T]$ ) can be categorized as ambiguous. When assessing the label for  $t = T - 2$  the condition for ambiguous demand checks if in the range of: the previous 3 months + current month + next  $T - (T - 2) = 2$  available months, at least  $3 + (T - (T - 2)) = 5$  have been labeled as experiencing a demand shock.

The labeling is retroactive as once data for the date  $t = T + 1$  is observed the classification of month  $T - 2$  -had it been classified as ambiguous- is potentially changed if it either satisfies the "complete" persistent supply or demand criterion. Otherwise it is labeled as transitory.

*NB1: with the sample denoted  $[1, T]$  we start the classification at date  $t = 1 + K$*

*NB2: the classification suggested above does not work for the parametric approach, as the latter simply does not have dummy variables for supply-demand classification but rather probabilities, which impedes the counting rule put in place.*

### 3. Results

#### 3.1. Results for France

FIGURE 1. Cross-correlations : demand component for France

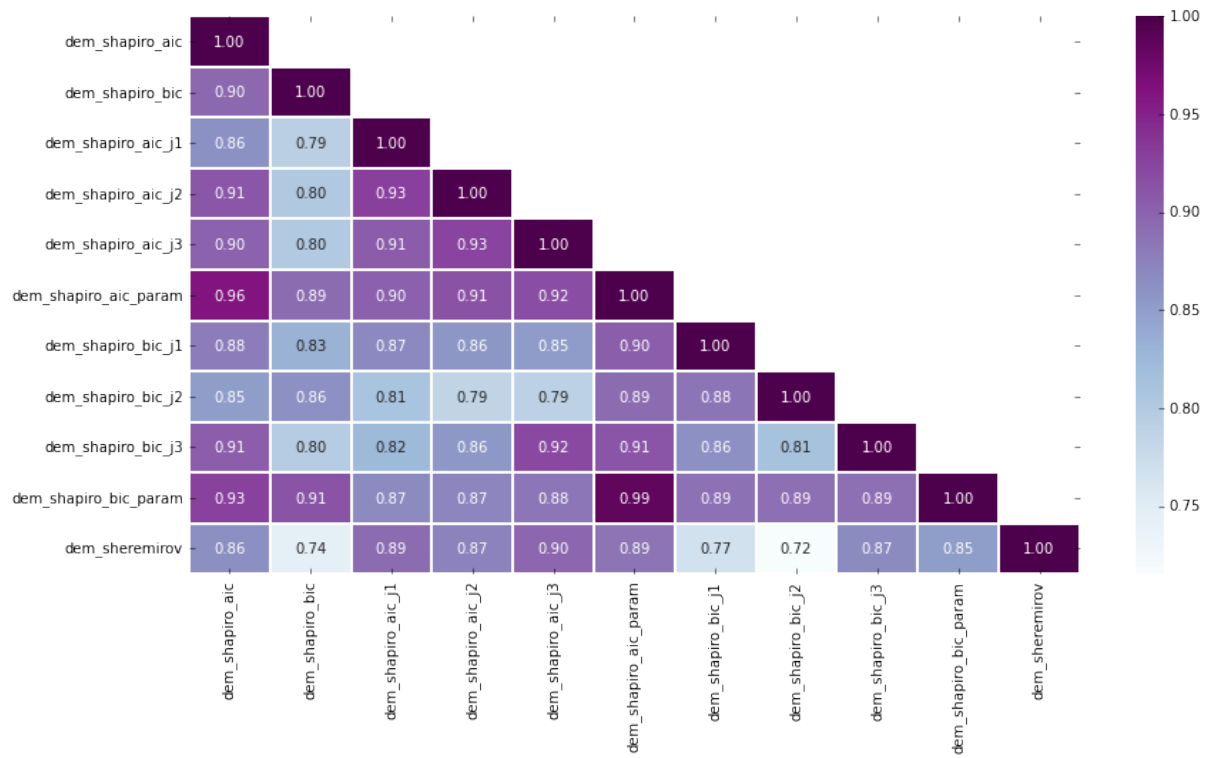


FIGURE 2. *Cross-correlations : supply component for France*

