

Utilisation des données Google Trends dans l'enquête de conjoncture Commerce de Détail de la Banque de France

- Contexte
 - Enquête CD
 - Livraison tardive des données FEVAD
- Google Trends & Modèle agrégé
 - Données :
 - Google Trends
 - Fiabilité des Google Trends
 - Jeux de variables initiaux
 - Indices conjoncturels quantitatifs
 - Modélisation :
 - Sélection de variables
 - Combinaison d'information
- Résultats
- Conclusion

L'Enquête Mensuelle de Conjoncture du Commerce de Détail

- <https://www.banque-france.fr/statistiques/chiffres-cles-france-et-etranger/enquetes-de-conjoncture/conjoncture-commerce-de-detail>
- Échantillon : 6 800 sources
- Grandeur économique : Taux de croissance annuelle du chiffre d'affaires
- Indicateurs quantitatifs conjoncturels :



Les CA FEVAD disponibles tardivement

CA FEVAD agrégés :

Électronique grand public

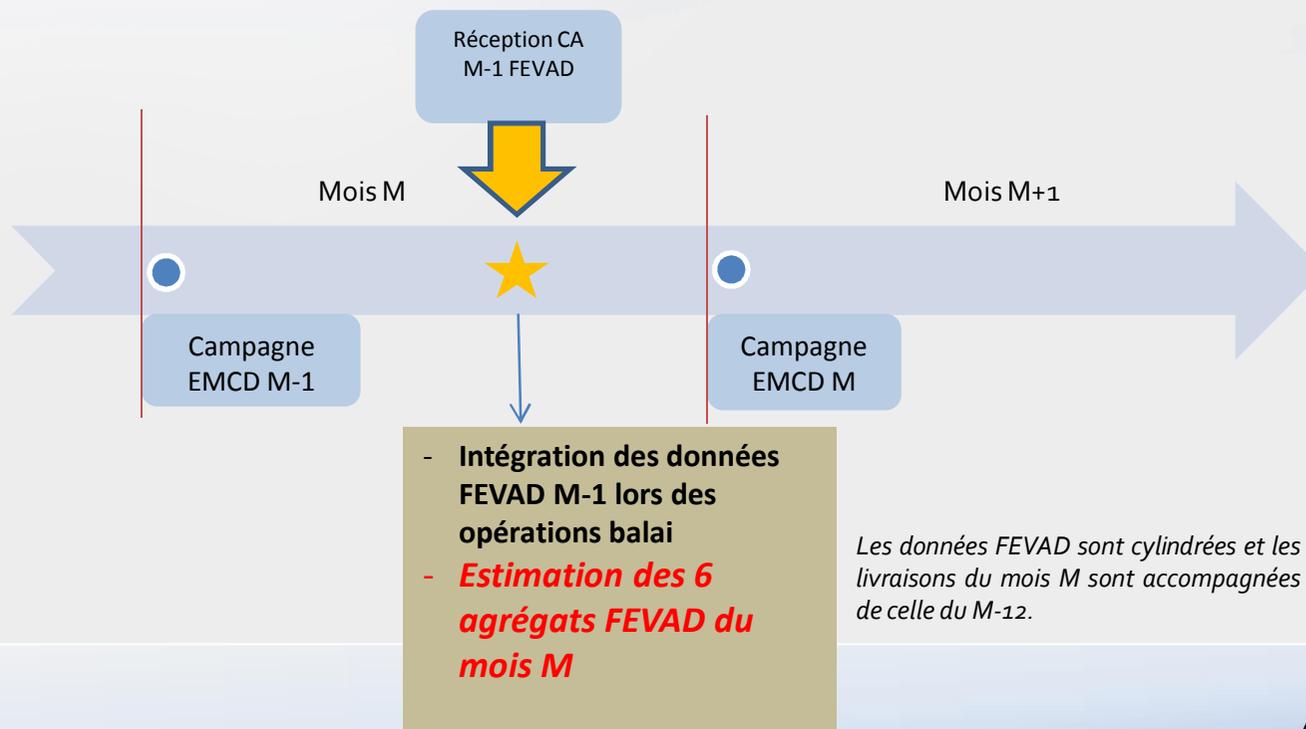
Électroménager

Total - Vente à distance

Habillement

Meuble

Chaussures



Données Google Trends

- Indices établis à partir d'échantillons
- Indices de valeurs entières [Kozicki & Hoffman (2004)]

Champs de la requête

Changements méthodologiques

Maximum établi à 100



Source Google Trends - FEVAD - Banque de France DGS SEEC

Google Trends

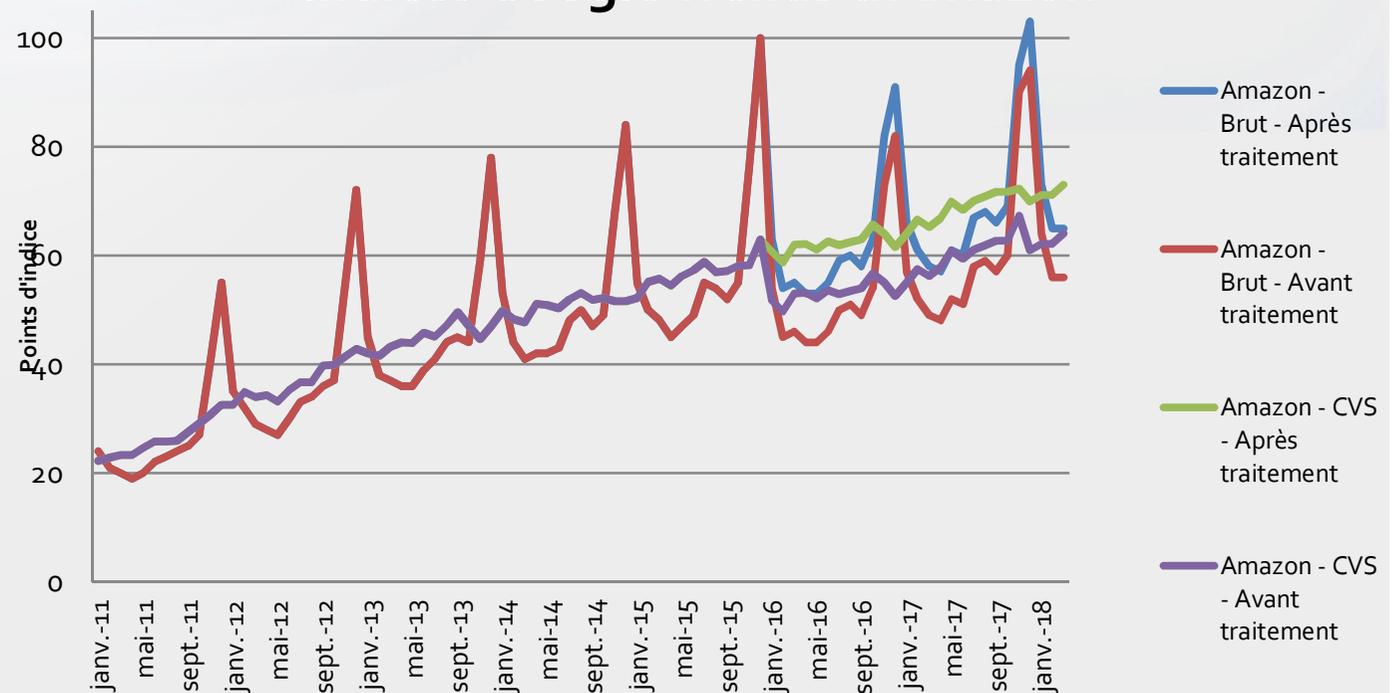
- Corrélations (%) entre 2 tirages :

Amazon	Cdiscount	Fnac	eBay	Carrefour
99,96	99,50	99,95	99,98	99,21

Source : Google Trends – FEVAD - Banque de France DGS SEEC

- Ruptures de séries :

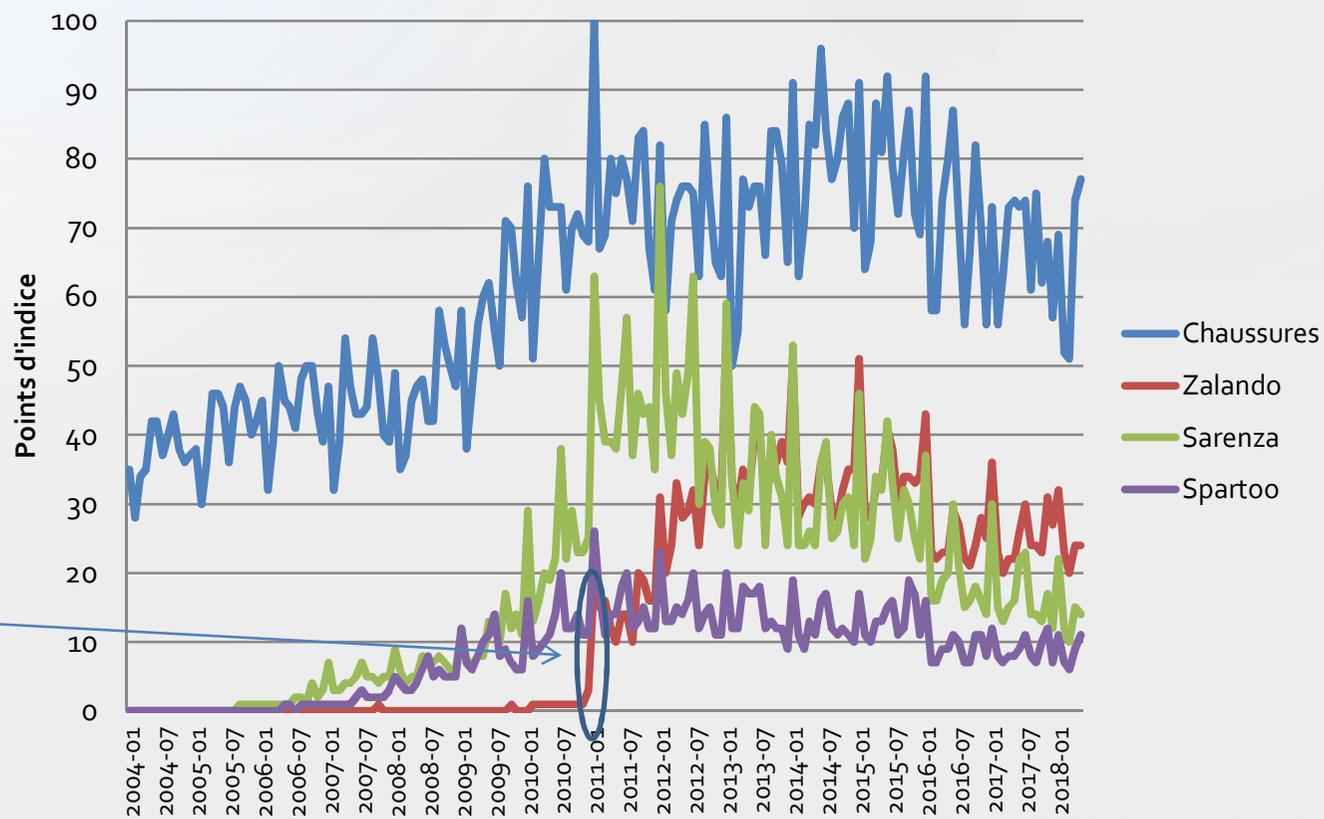
Indices Google Trends d'Amazon



Source : Google Trends – FEVAD - Banque de France DGS SEEC

Google Trends

Indices Google Trends liés à la chaussure

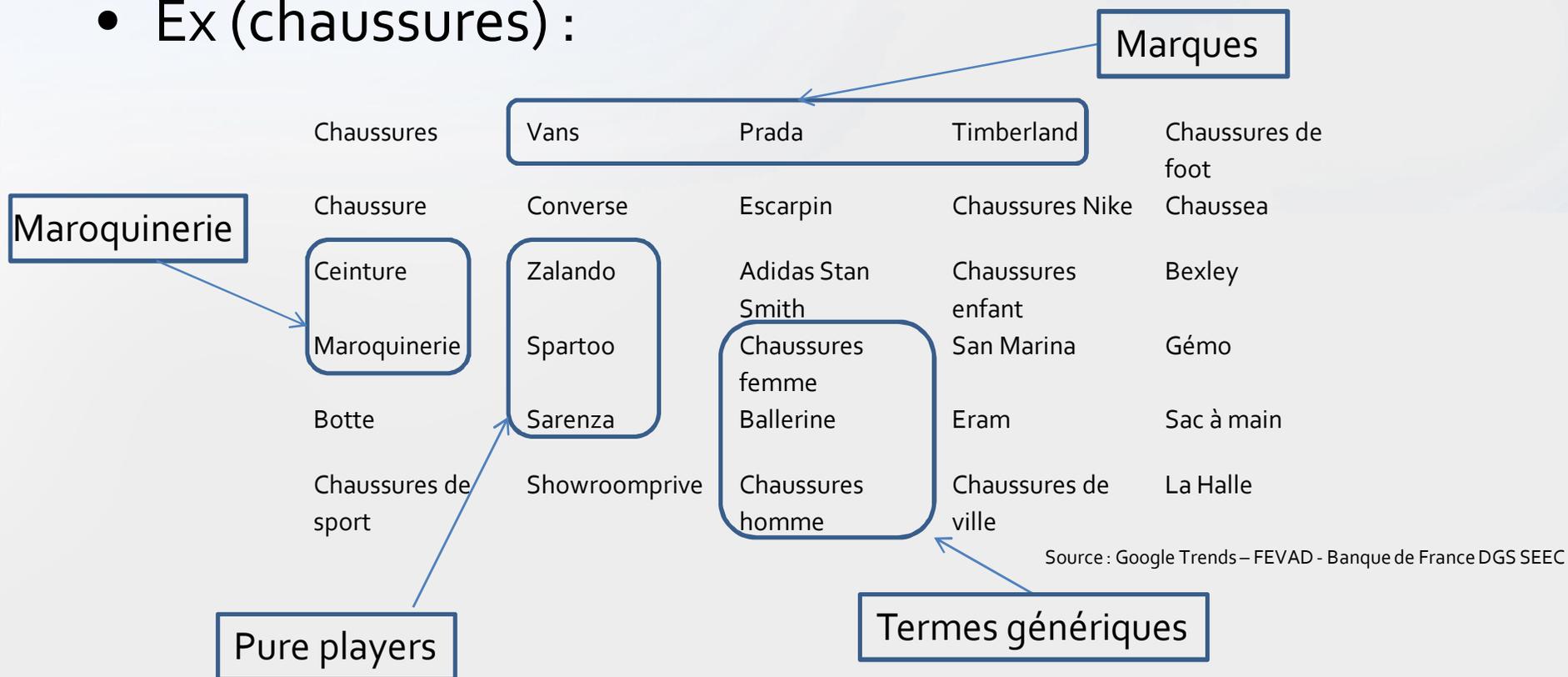


Source : Google Trends – FEVAD - Banque de France DGS SEEC

Nouvel acteur

Jeux de variables initiaux – Google Trends

- Variables hétérogènes
- Ex (chaussures) :



Variables CD

- Indices quantitatifs conjoncturels utilisés :

	Petit commerce traditionnel	Grande surface spécialiste et succursaliste	Hyper et supermarchés	Grand magasin et magasin populaire	Ensemble des ventes physiques
Total Produits industriels hors automobiles	0	0	0	0	1
Chaussures	1	1	0	0	0
EGP	1	1	1	0	0
Électroménager	1	1	1	0	0
Meubles	1	1	1	0	0
Habillement	1	1	1	1	0

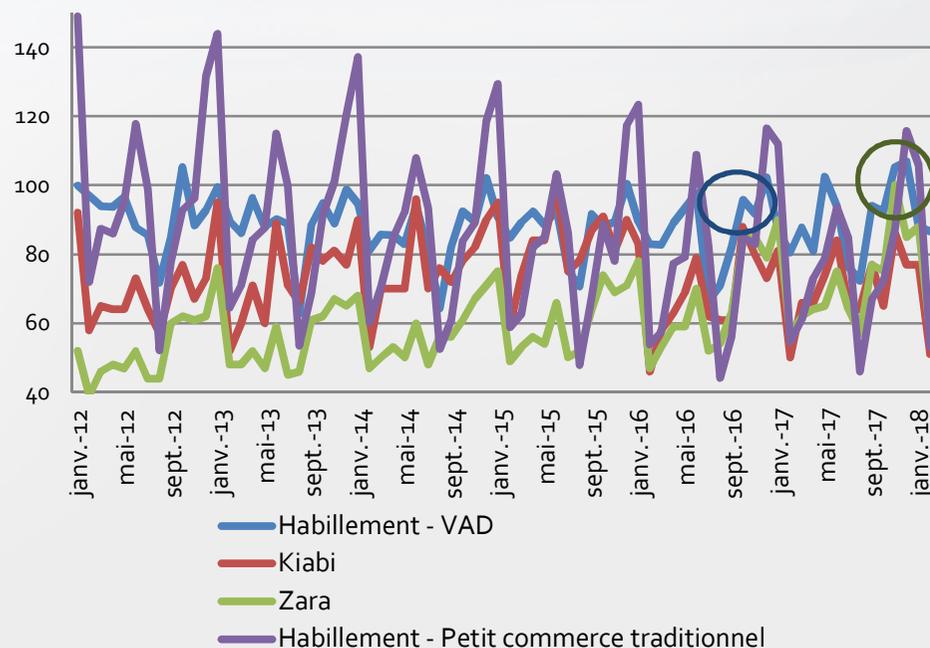
Source : Google Trends – FEVAD - Banque de France DGS SEEC

- Pour le total : indices VAD des produits (prolongés *via* SARIMA)

Données

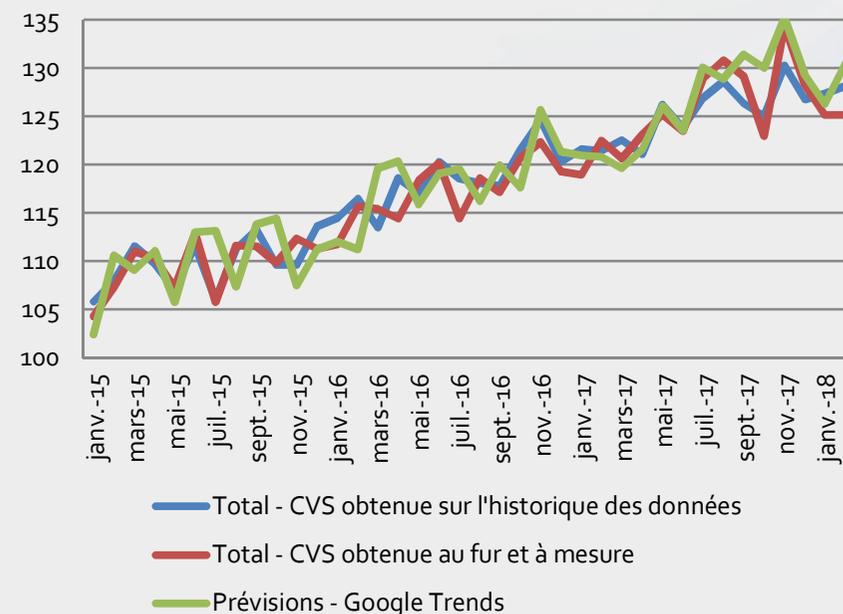
- Données brutes différenciées => CVS:
- Instabilité (derniers points) & Changements de saisonnalité

Indices dans l'habillement



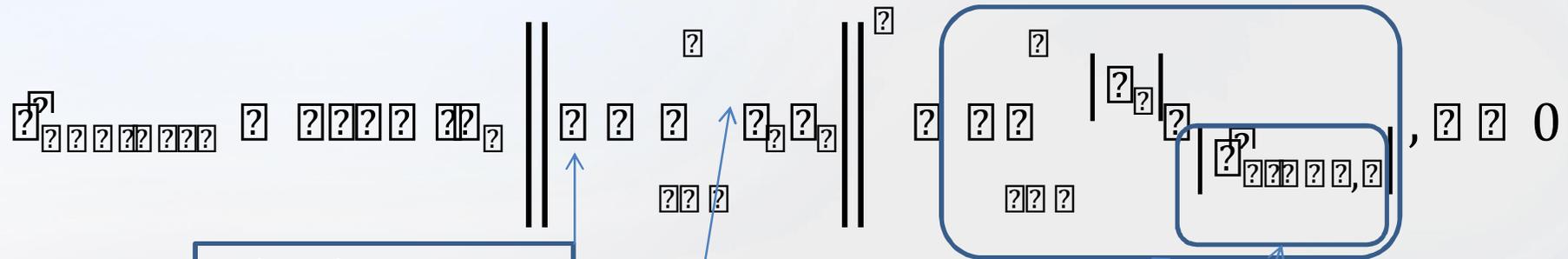
Source : Google Trends – FEVAD - Banque de France DGS SEEC

Stabilité des CVS



Source : Google Trends – FEVAD - Banque de France DGS SEEC

” Choix du modèle : Adaptive Lasso



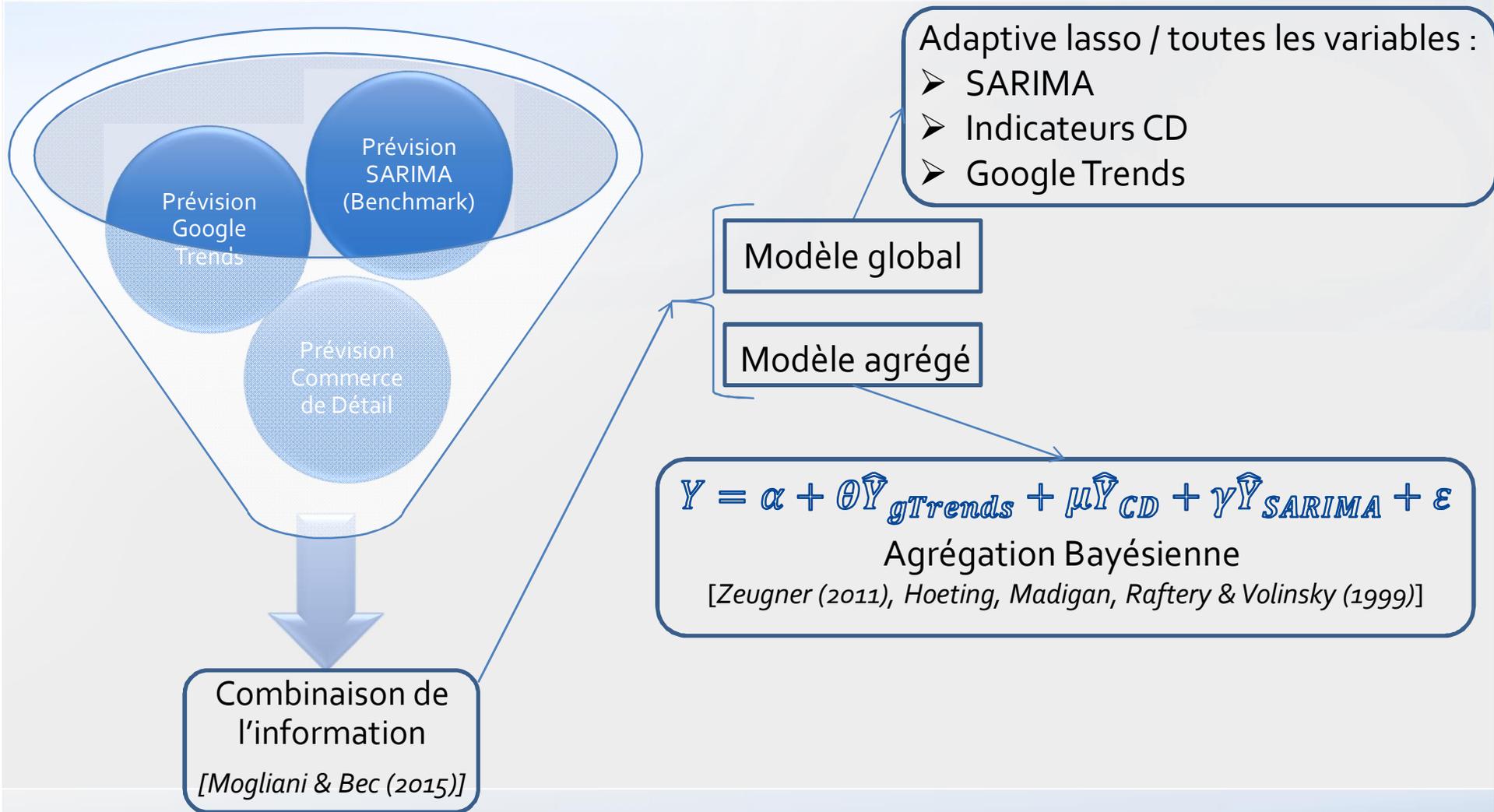
Indice de CA FEVAD

Variables exogènes : Google Trends
Ou indices quantitatifs conjoncturels
+ tendance & composante AR (SARIMA)

LASSO [Tibshirani, 1996] : régression pénalisée

Adaptive LASSO [Zou, 2006] : pénalités individuelles
(ici déterminées par l'estimateur Ridge)

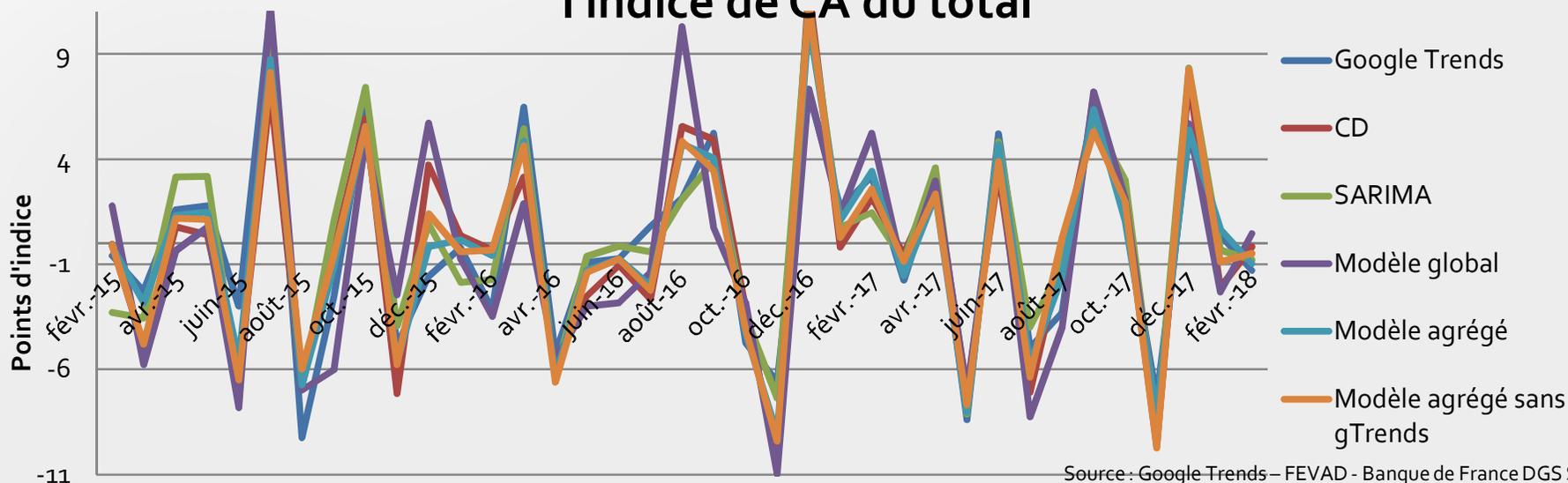
Modèles



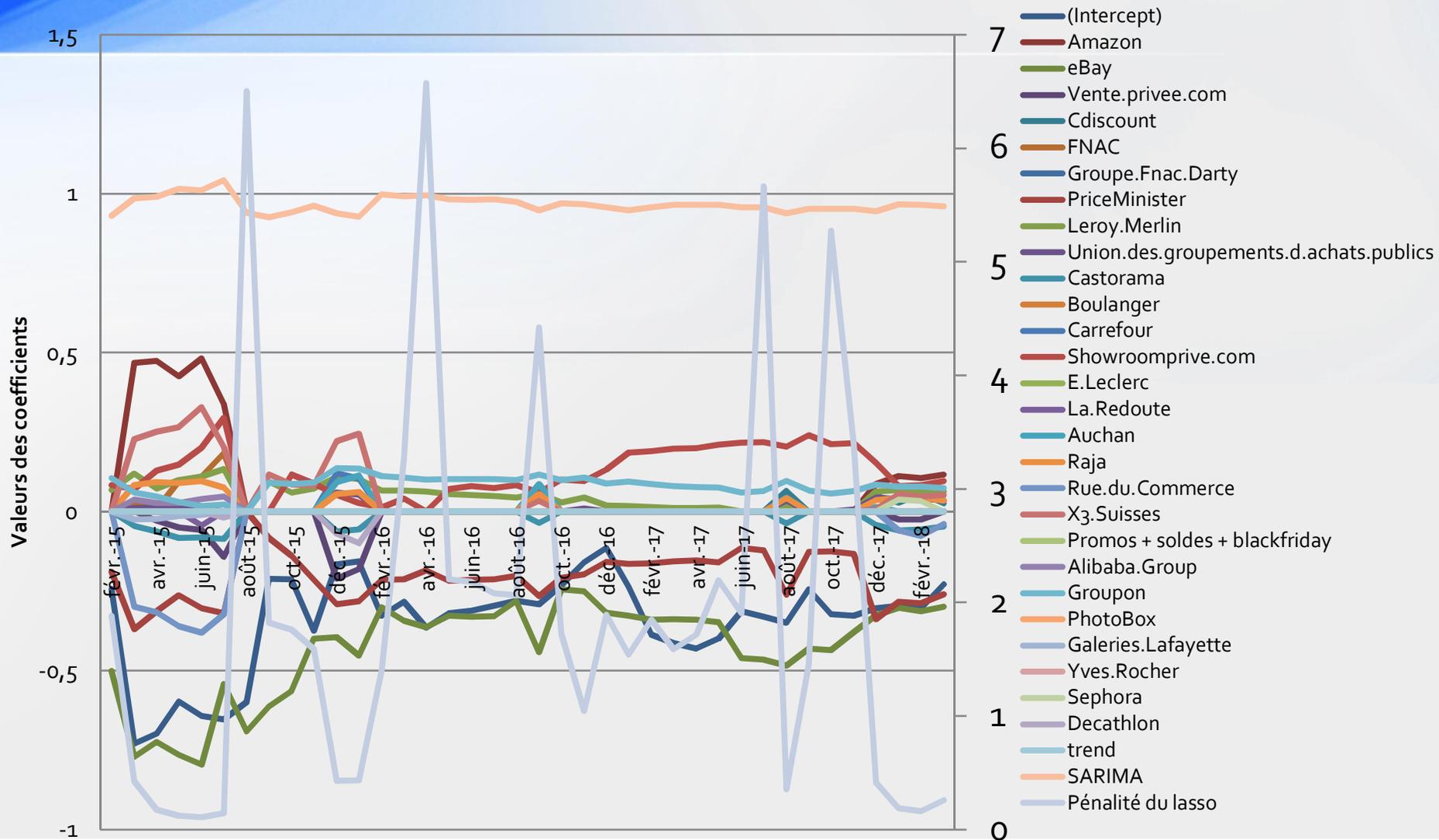
Total	Google Trends	CD	SARIMA	Modèle global	Modèle agrégé	Modèle agrégé sans gTrends
Moyenne des RMSE	3,3	3,7	4,2	2,3	2,6	2,8
RMSFE	4,8	5,2	5,0	5,5	4,8	5,0
Moyenne des erreurs de prévision absolues	3,9	4,0	3,9	4,5	3,8	3,9

Source : Google Trends – FEVAD - Banque de France DGS SEEC

Erreurs de prévision des différents modèles - estimation de l'indice de CA du total



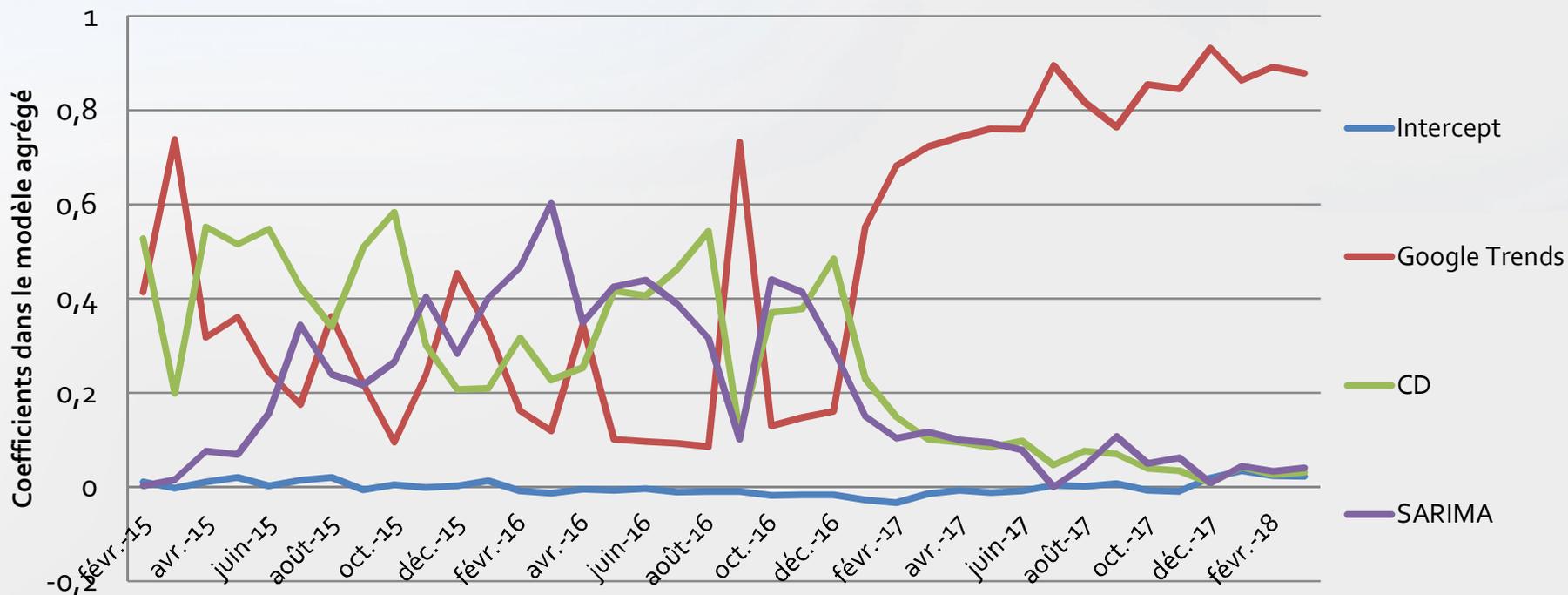
Evolution des coefficients et de la pénalité LASSO du modèle Google Trends – Total



Source : Google Trends – FEVAD - Banque de France DGS SEEC

Pondération des modèles dans l'agrégation – Total

Pondération dans l'agrégation - Total



Source : Google Trends – FEVAD - Banque de France DGS SEEC

Résultats - produits

RMSFE	Google Trends	CD	SARIMA	Modèle global	Modèle agrégé
Chaussures	13,2	12,7	13,6	13,8	13,4
Meubles	11,9	12,3	12,0	13,2	11,8
Électroménager	11,7	10,4	10,2	12,3	11,2
EGP	15,5	15,3	16,4	11,5	13,1
Habillement	9,8	10,1	9,2	15,2	9,7

Source : Google Trends – FEVAD - Banque de France DGS SEEC

- “ Opacité de la construction des données Google :
 - Méthode de comptabilisation non connue
 - Évolutions du moteur de recherche

- “ Phénomènes mouvants :
 - Acteurs changeants
 - Modes de consultation internet préalable à un achat
 - Nécessité de réviser (annuellement) les modèles

- “ Résultats mitigés : positifs sur le total et encourageants sur certains produits
- “ Information multidimensionnelle
- “ Utilisation en production

Annexes

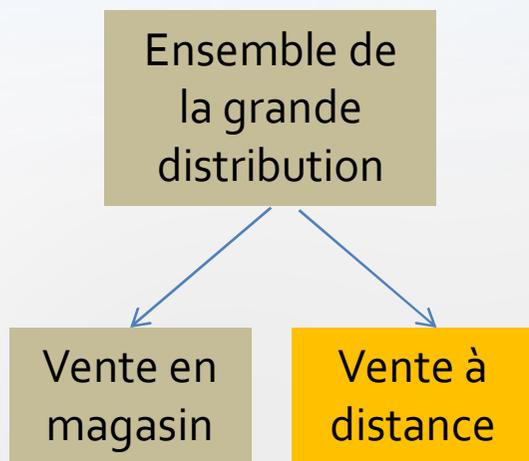
Google Trends

Total	Habillement	EGP	Chaussures	Électroménager	Meubles
Amazon	Vertbaudet	iPhone	Chaussures	Clubic	But
eBay	Kiabi	Apple	Chaussure	Boulangier	Legallais
Vente privée	H&M	Cdiscount	Ceinture	Cdiscount	Cuisine
Cdiscount	C&A	PC Gaming	Maroquinerie	Four	Raja
Fnac	Jules	iPad	Botte	Réfrigérateur	Staples
Groupe Fnac Darty	Linge de maison	Groupe Samsung	Chaussures de sport	Machine à laver	Leroy Merlin
PriceMinister	Zara	FNAC	Vans	Darty	Castorama
Leroy Merlin	Costume	Télévision	Converse	Bosch	Conforama
UGAP	Culotte	Boulangier	Zalando	Electrolux	Vega
Castorama	3 Suisses	Sony	Spartoo	Conforama	Bureau
Boulangier	Devred	LDLC Pro	Sarenza	Amazon	Meuble
Carrefour	Robe	Amazon	Showroomprive	Cuisinière	Roche Bobois
Showroomprive	Etam	Phillips	Prada	Électro-dépôt	Armoire + étagère
E. Leclerc	La Redoute	LG Group	Escarpin	Brandt	Ikea
La Redoute	Jeans + Chino + Pantalon	Samsung Electronics	Adidas Stan Smith	Four à micro-ondes	Couteaux
Auchan	Manteau + Blouson	Téléphone + smartphone	Chaussures femme	Groupe Fnac Darty	Maisons du Monde
Raja	Veste	Darty	Ballerine	Aspirateur	Cinna
Rue du commerce	Vêtement femme	Tablette tactile	Chaussures homme	Whirlpool Corporation	Meuble en bois
3 Suisses	ASOS	Enceinte	Timberland	Mistergooddeal	Roset
Promos + soldes + blackfriday	Maisons du Monde	Appareil photographique reflex numérique	Chaussures Nike	GrosBill	Buffet + commode + vaisselier
Alibaba	Lingerie	Ordinateur portable +PC	Chaussures enfant	Pulsat	Table + chaise + canapé
Groupon	Jennyfer	Bose	San Marina	Ubaldi	Fauteuil
PhotoBox	Vêtement	JBL	Eram	But	
Galeries Lafayette	Galeries Lafayette	Groupe Fnac Darty	Chaussures de ville		
Yves Rocher	Bonobo	Barre de son	J.M Weston		
Sephora	Brandalley	Appareil photo	Chaussea		
Decathlon	Camaieu	Marshall	Bexley		
	Showroomprive		Gémo		
	Vente Privée		Sac à main		
	Rideau		La Halle		
	Blanche Porte		Chaussures de foot		
	Drap				
	Coussin				
	Homemaison				
	Sous-vêtement				
	La Halle				
	Decathlon				

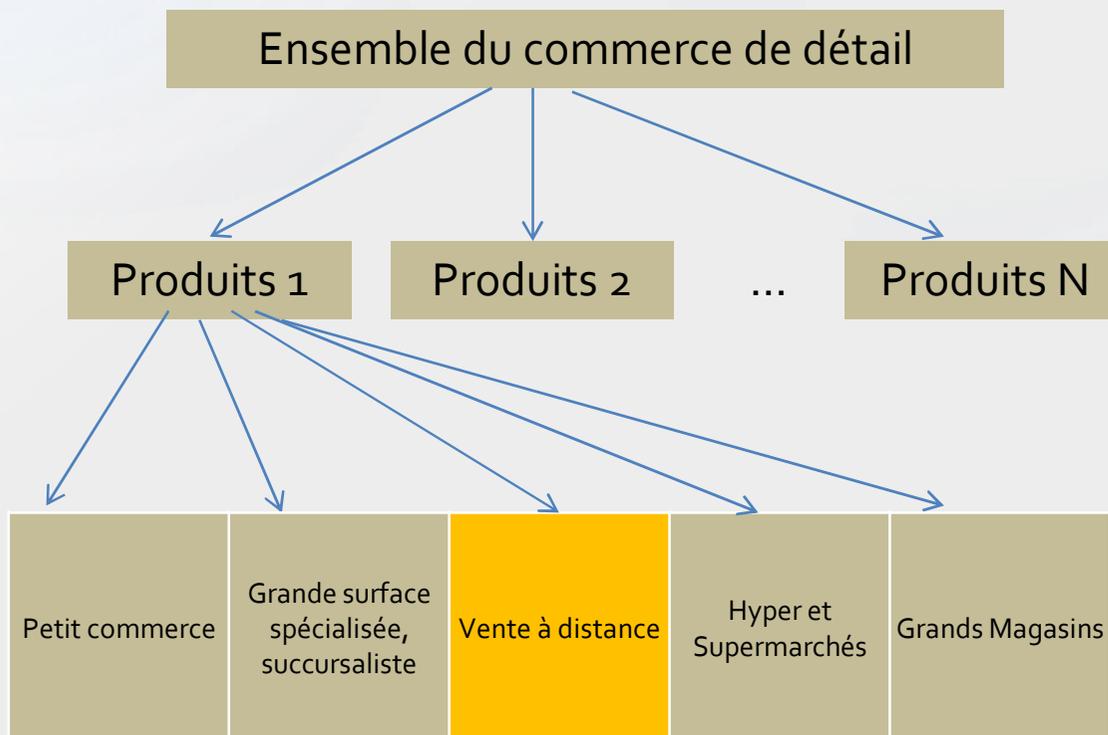
Source : Google Trends – FEVAD - Banque de France DGS SEEC

Le e-commerce dans l'enquête, une strate parmi les autres

Déclinaison par type de distribution



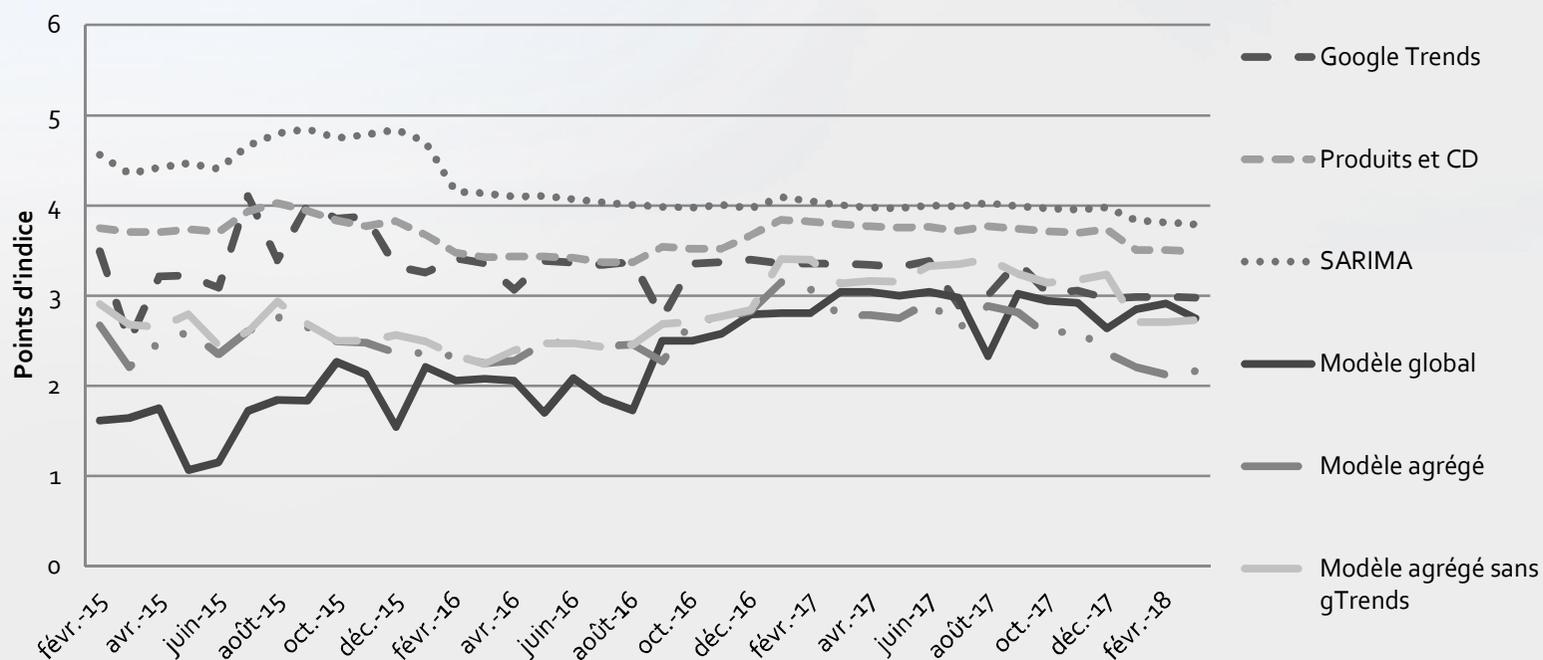
Déclinaison par produit puis par type de distribution



Source : Google Trends – FEVAD - Banque de France DGS SEEC

RMSE - Total

RMSE des modèles



Source : Google Trends – FEVAD - Banque de France DGS SEEC

Variables - Total

	Moyenne	Minimum	Maximum
Amazon	0,07	0,00	0,48
eBay	-0,43	-0,80	-0,24
Vente.privee.com	-0,02	-0,21	0,00
Cdiscount	0,01	0,00	0,07
FNAC	0,01	0,00	0,18
Groupe.Fnac.Darty	0,00	0,00	0,00
PriceMinister	-0,21	-0,37	0,00
Leroy.Merlin	0,05	0,00	0,13
Union.des.groupements.d.achats.publics	0,01	0,00	0,06
Castorama	-0,02	-0,09	0,00
Boulangier	0,00	0,00	0,00
Carrefour	0,01	0,00	0,12
Showroomprive.com	0,12	0,00	0,29
E.Leclerc	0,00	0,00	0,00
La.Redoute	0,00	-0,04	0,00
Auchan	0,01	0,00	0,11
Raja	0,02	0,00	0,10
Rue.du.Commerce	-0,05	-0,38	0,00
X3.Suisses	0,06	0,00	0,33
Promos + soldes + blackfriday	0,00	0,00	0,00
Alibaba.Group	0,00	0,00	0,05
Groupon	0,08	0,00	0,14
PhotoBox	0,00	0,00	0,00
Galleries.Lafayette	0,00	-0,02	0,00
Yves.Rocher	0,00	0,00	0,00
Sephora	0,00	0,00	0,04
Decathlon	0,00	-0,10	0,00
trend	0,00	0,00	0,00
SARIMA	0,97	0,93	23 1,04

Variables retenues - Produits

Chaussures	Spartoo (38)	Sarenza (36)	Converse (36)	Chaussures de ville (32)	Chaussures de foot (28)
Meubles	Cinna (38)	Roset (33)	Meuble en bois (32)	Buffet + commode + vaisselier (31)	IKEA (26)
Électroménager	Machine à laver le linge (37)	Four (28)	Cuisinière (28)	Conforama (26)	GrosBill (24)
EGP	Appareil photo reflex numérique (37)	Télévision (35)	JBL (35)	Sony (27)	Samsung Electronics (23)
Habillement	Costume (34)	Décoration (29)	Jennyfer (28)	Lingerie (27)	3 Suisses (25) et Vêtements femme (25)

Source : Google Trends – FEVAD - Banque de France DGS SEEC

Agrégation Bayésienne

- $\hat{\theta} = \sum_{k=0}^7 P(M_k|D)\hat{\theta}_k$
- Avec :
 - D les données
 - M_k le modèle k
 - $\hat{\theta}_k$ le paramètre obtenu dans le modèle M_k
 - $P(M_k|D)$ la probabilité que le modèle M_k soit le bon