Links between Google Search and Payment Digitization: What assumptions matter?

Sanvi Avouyi-Dovi [1], Louis-Alexandre Bayol [2], Lucas Devigne [2], Ambre Geraut [3] and Christian Pfister [4]

Preliminary version

[1] ICN Business School
[2] Banque de France
[3] CRED, Inseec Grande École
[4] Université d'Orléans

Disclaimer : the views expressed in this paper are those of the authors and should not be interpreted as reflecting the views of the ICN Business School, Banque de France, CRED – INSEEC and Université d'Orléans. Any errors or omissions are the responsibility of the authors.

2026 Carpathian Stata Conference

July 27th, 2026 Cluj-Napoca

What is Google Search Volumes data?

- Google Search Volumes data, commonly referred to as Google Trends indices, capture the search interest of users in real time.
- GSV provides a window into consumer attention, curiosity, and information-seeking behavior related to payment methods.
- In other words, by analyzing search queries for terms such as “card payment”, "contactless payment," "Apple Pay," "mobile wallet," and "digital payment," we can assess a granular measure of the digitization process that replaces or complements traditional statistics.

Google Trends

Why do we need a measure of payment digitization?

What official statistics offer?

- Annual, half-yearly transactions records and central bank statistics, biennial consumer surveys (e.g. ECB's SPACE).
- Indicators focus on a single payment method or technology, or rely on ad hoc weighting schemes : each captures only one facet of the transformation.

What is missing?

- Publication delay, large granularity, large coverage, etc.
- A multi-dimensional vision
- Failure to take into account certain features or assumptions

Google Search Volumes as a window into payment behavior

What can search data tell us about digitization?

- Millions of real-time queries reveal consumer attention, curiosity and information-seeking behavior.
- Search interest in "contactless payment," "Apple Pay" or "mobile wallet" mirrors actual adoption decisions.
- Search-based indicators complement, and sometimes anticipate, official statistics.

Google Search Volumes as Digitization index

- To know if search data can tell us more about digitization.
- To build the indicator best captures means of payment Digitization.
- To summarize the information from GSV into a single indicator by testing several hypothesis and investigate their consistency.
- To scrutinize the understanding of payment Digitization and compare with existing Digitization indicator.

Research question and contribution

Can Google search data capture, and even anticipate, the digitization of payment methods, and which modeling hypothesis best captures this dynamic?

1) Collecting a multidimensional database of payment-digitization proxies from GSV
2) Comparing the indicators drawn from three competing hypotheses on the nature of the underlying process, non-linearity (Kernel-PCA), sequential dependence (Long Short-Term Memory, LSTM), temporal dynamics (Dynamic Factor Model, DFM).
3) Asses the empirical relevance of the indicators by comparing their dynamics with those of economic, financial and monetary aggregates (next step)

The promise and the limits of search-based indicators

- Search-based indicators significantly strengthen short-term forecasting performance for unemployment, consumption and GDP Askitas and Zimmermann (2009); Schmidt and Vosen (2010)
- Synthetic indicators built from search volumes successfully nowcast inflation and consumption once combined with appropriate models Bikourne et al. (2025); Tenorio et al. (2025)
- Yet these indicators face well-known challenges: limited transparency of search algorithms, sampling noise, and incomplete population coverage Bikourne et al. (2025); Tenorio et al. (2025)
- Composite indices built via Principal Component Analysis summarizing multiple series, offer a broader view Stock and Watson (2002); D’Amuri and Marcucci (2012); Götz and Knetsch (2019); Bikourne et al. (2025).

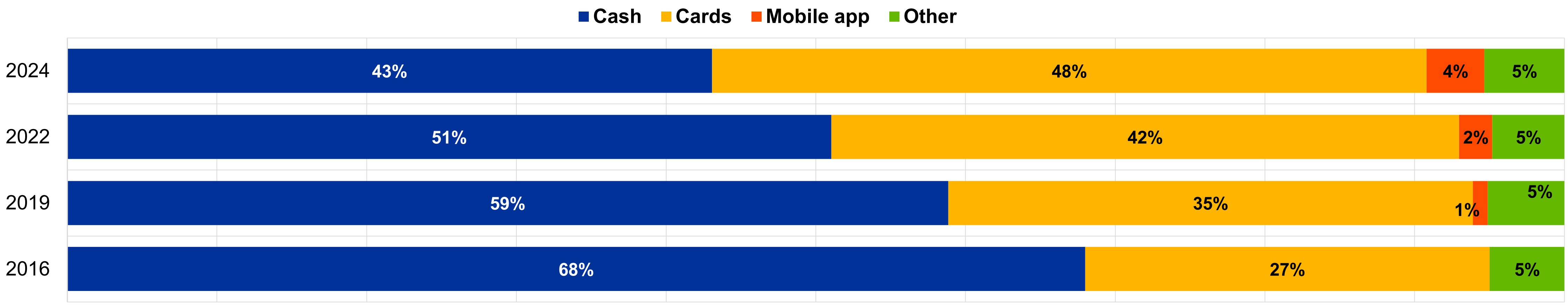
The promise and the limits of search-based indicators

- Google queries provide fast, reliable signals for unemployment and consumption, often outperforming survey-based measures Askitas and Zimmermann (2009), Choi and H. Varian (2012), Schmidt and Vosen (2010)
- Factor-based dimensionality reduction methods provide the methodological foundations for high-dimensional forecasting with search data Stock and Watson (2002)
- Internet search volumes rank among the most promising sources for macroeconomic nowcasting, owing to their timeliness and low collection cost Tenorio et al. (2025)
- Search spikes capture investor attention and predict short-term price movements and reversals Da, Engelberg, and Gao (2011)
- Regional search interest has been used as an instrument to identify the causal effects of technology adoption Diebold (2025)
- Large Language Model approach combined with newspaper article data has been used to build a multidimensional indicator of Digitization of Payment, Avouyi-Dovi et al, (2026)

Learnings from Study on Payment Attitude (SPACE) in France

- The use of cashless payment for retail transactions has increased since 2016.
- Share of cashless payment at the POS in volume (card and mobile) increased by 25 pp from 27% in 2016 to 52% in 2024 (t point of sale has decreased by 25 percentage points (pp) from 68% in 2016 to 43% in 2024.
- SPACE surveys confirm increasing trend in payment method digitization.

Chart 1 : Share of payment instruments used at the POS

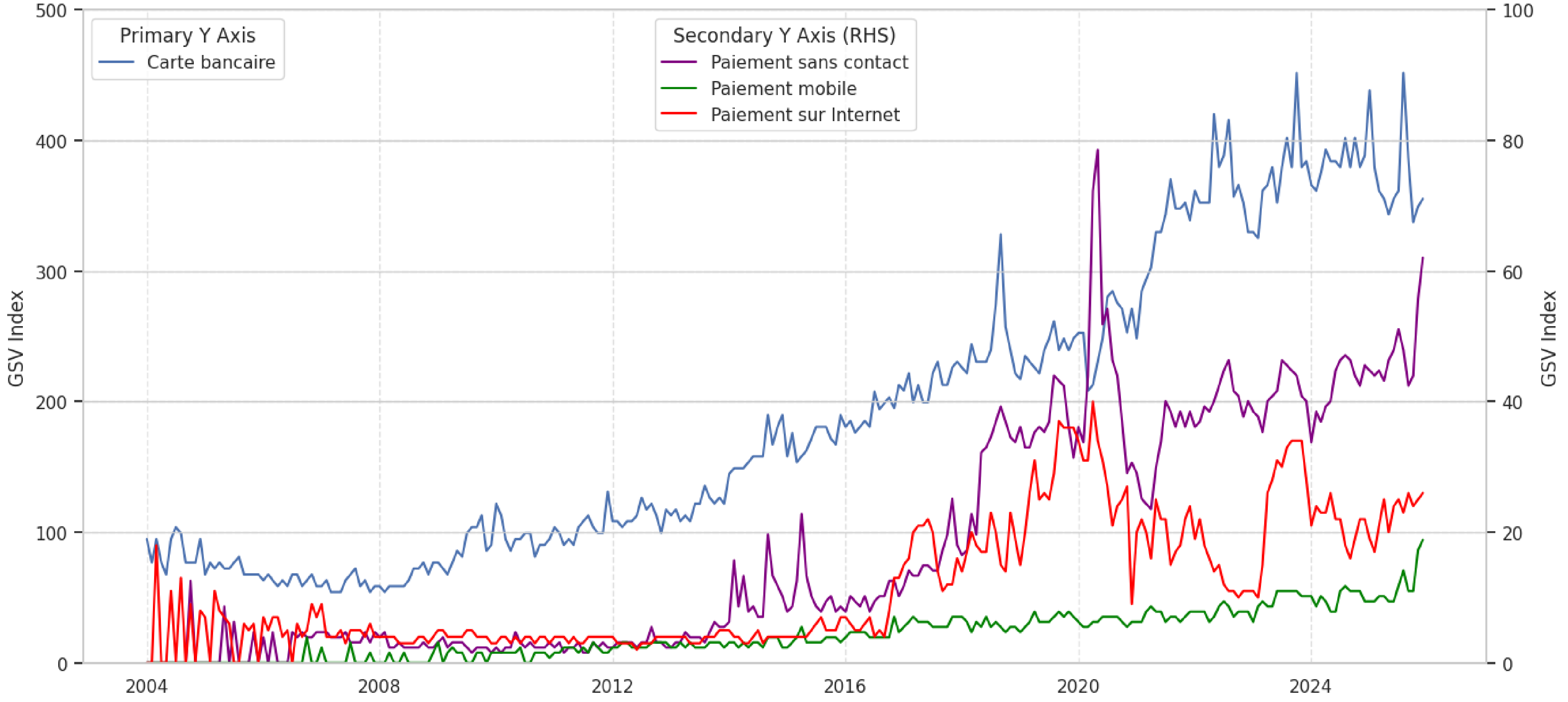


Source: ECB SPACE.

Payment methods in France

- Payment card is leading the payment digitization in France.

Chart 2: Google Search Volumes (GSV) for card, cash, and mobile payments in France (2004-2025)

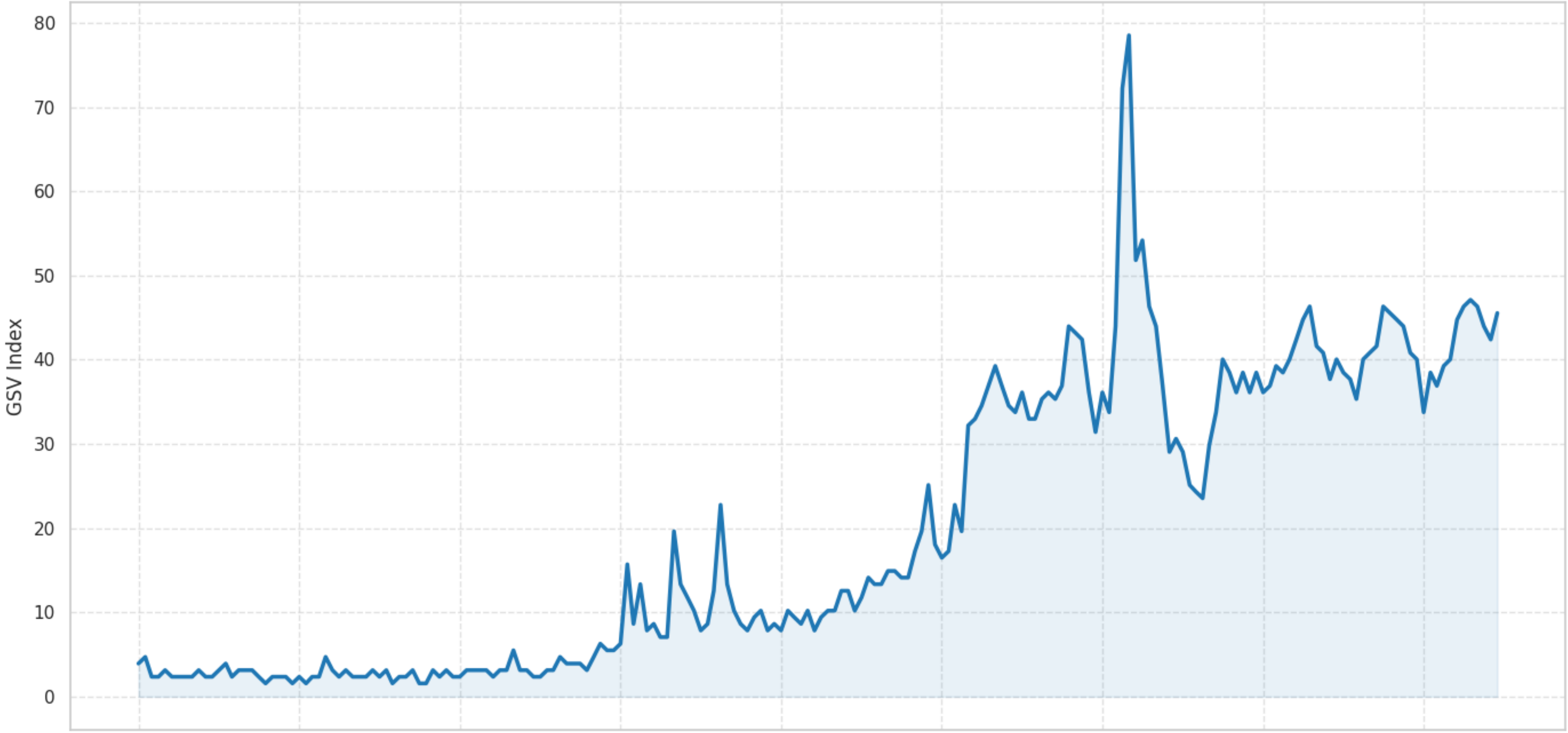


Source: author’s calculations

Payment methods in France

- Google searches for contactless payment peaked in 2020, driven by the COVID-19 pandemic.
- Reflecting increased the contactless card payment cap in France in In May 202 → further boosting public interest and adoption.

Chart 3: GSV for Contactless Payments in France (2008-2025)

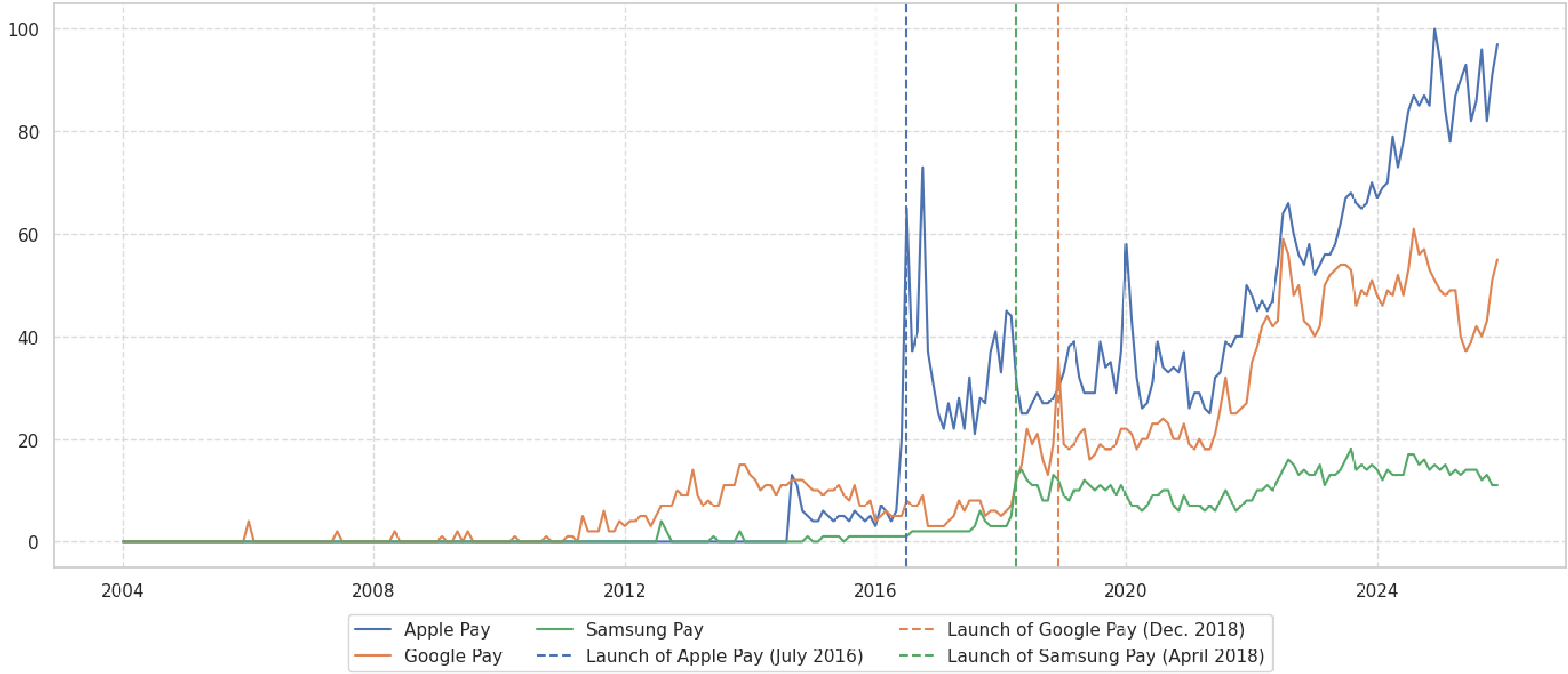


Source: author's calculations

The Rise of Mobile App Payments

- Mobile payment searches were initially driven by product launches but gradually became part of a broader and lasting adoption trend

Chart 4: GSV on mobile payment applications



Sources: GSV and authors' calculations.

Google Search Volumes (GSV)

- GSV capture search interest based on actual user behavior
- Monthly GSV indices normalized between 0 and 100
- Keywords related to payment methods and digitalization in France
- GSV provides real-time information on consumer interest and adoption of payment technologies (Eichenauer et al. ,2021)

Querying Google Trends

- Selection of research topics linked to payment digitization
- All GSV series are normalized using a common benchmark topic “Online payment to ensure temporal comparability
- Online payment refers to the topic of payment and digitization of payment that insure consistency in search intent.

Table 1: List of keywords, topics and applications in Google research related to Digitization of means of payment

Query 1	Query 2	Query 3
Pivot topic of research • *Paiement sur Internet* (“Paiement sur Internet”) **Application in research** • Apple Pay • Google Pay • Samsung Pay • Amazon Pay • Lydia	**Pivot topic of research** • Online payment (“Paiement sur Internet”) **Topics of research** • Mobile payment (“Paiement mobile”) • Contactless payment (“Paiement sans contact”) • Wallet (“Portefeuille numérique”) • Card (“Cartes”) - Apple application) • Google Wallet	**Pivot topic of research** • Online payment (“*Paiement sur Internet”)* **Research terms** • Bank card (“carte bancaire”) **Topics of research** • Bank card (“Carte bancaire”)

Sources: Authors’ calculation.

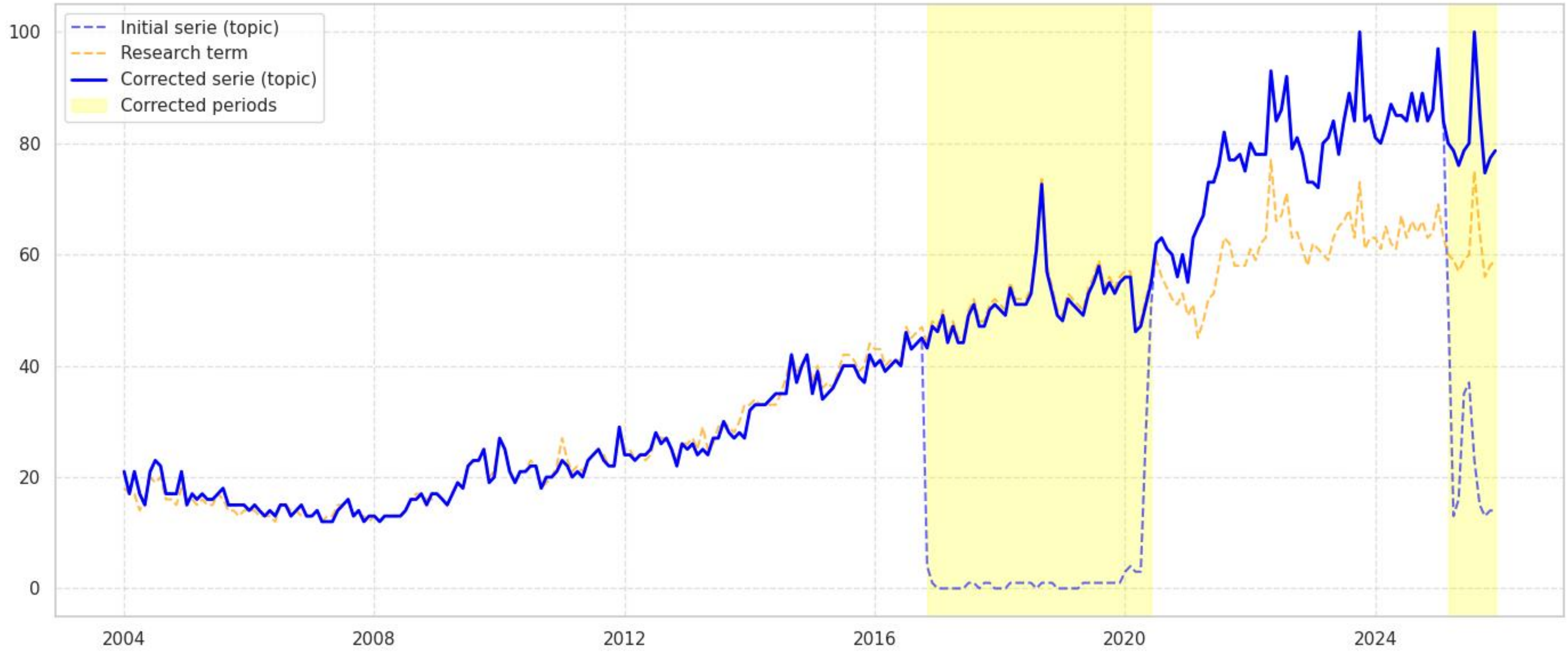
From dependency series to a common factor

- Google search volume indicators are standardized (mean-centered, unit variance) at a monthly frequency, January 2004 to December 2025.
- Extract the long-term co-trending component.
- We ask which modeling hypothesis best summarizes the common information they share, and build the Digital Payment Index (GSV-DPI) as the resulting latent factor.

Correction of GSV anomalies

- We combine the broad semantic coverage of Google Topics with the stability of exact search terms to patch the topic series and provides a smooth.
- The reliable trend shows sustained growth in digital and card payment search interest in France from 2004 through 2025.

Chart 5: Applied correction on "bank card" topic with research term



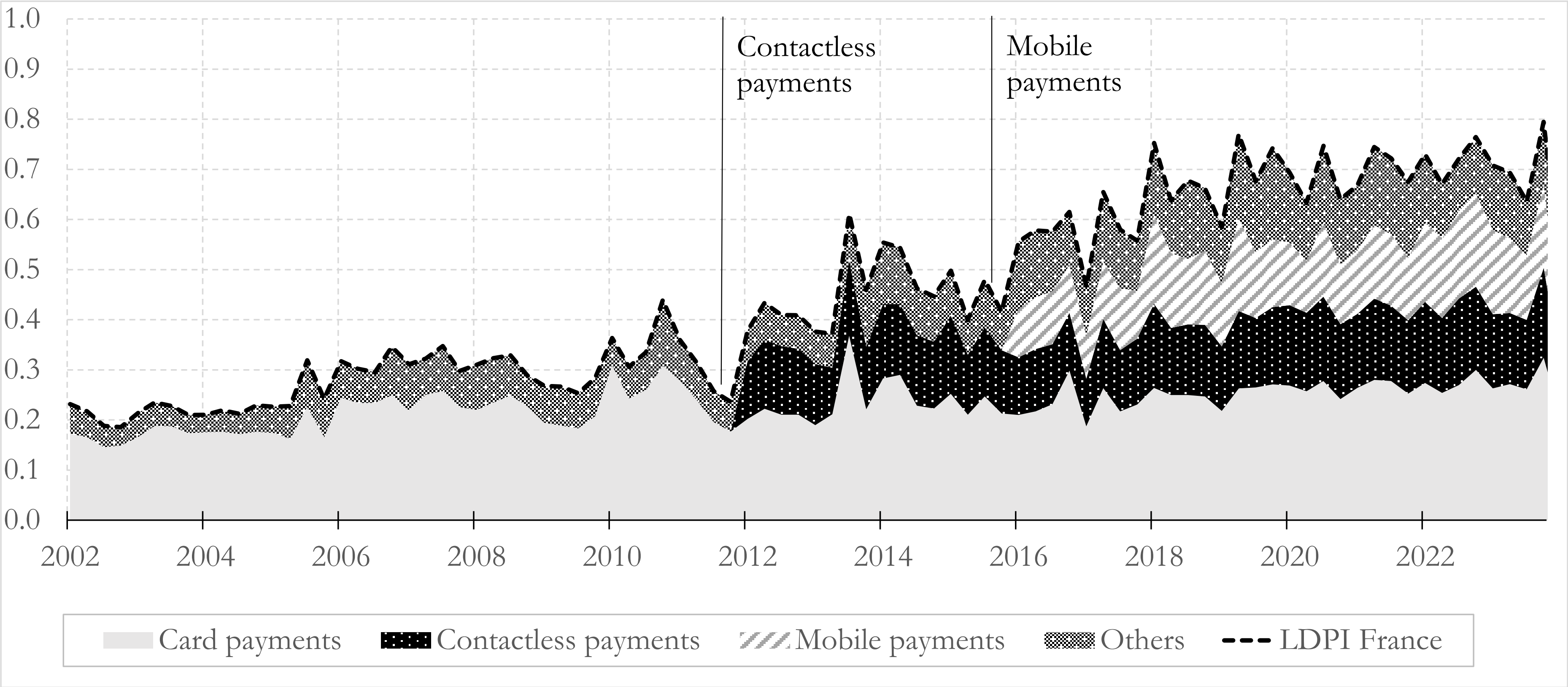
Sources: Authors' calculation.

LLM-based Digitization of Payments Index (LDPI)

Digitization of payments in France: Lessons from large language models (LLM)

- As an external validation, we compare the GSV-DPI to an independent index built from a Large Language Model approach
- Digitization of Payments indicator in France (see Avouyi-Dovi et al, 2026, The Quarterly Review of Economics and Finance)

Chart 6: The LLM-based Digitization of Payments Index (LDPI) digitization indicators for France



Notes: The sum of the sub-classes corresponds to the LDPI. Contactless and mobile payment technologies were introduced in France in 2012 and 2016, respectively. The index has been normalized using the min-max scaling method.
Sources: Authors' calculation.

Three hypotheses testing

a) **Non-linearity**: The relationship between search series may be non-linear, requiring an approach like a Kernel-based extension of for factor extraction.

b) **Temporal dynamics:** Dynamic Factor Model is appropriate to describe the latent factors which may evolve smoothly over time.

c) **Sequential dependence:** the signal may depend on the order and memory of past search behavior, not just contemporaneous correlations. The LSTM autoencoder approach can help to face this characteristic.

Kernel PCA index

- Kernel PCA transforms correlated variables into uncorrelated components in a higher-dimensional feature space, ordered by the variance each one explains.
- This captures non-linear relationships among search indicators that a linear approach cannot identify.
- First component captures the maximum possible variance and becomes the basis of the GSV-DPI.

DFM model

- The Dynamic Factor Model (DFM) is a statistical and econometric framework designed to extract a few unobserved (latent) component that drive the common movements across a large set of observed time-series variables.
- The Dynamic Factor Model extends the static factor framework by explicitly modeling the temporal evolution of the latent digitization factor while standard Static Factor Analysis (PCA) only captures cross-sectional correlations at a single point in time.
- It models how these underlying factors evolve over time, interact with past states, and react to dynamic shocks.
- It separates trend from cyclical movement and produces a smoothed latent factor via the Kalman filter.

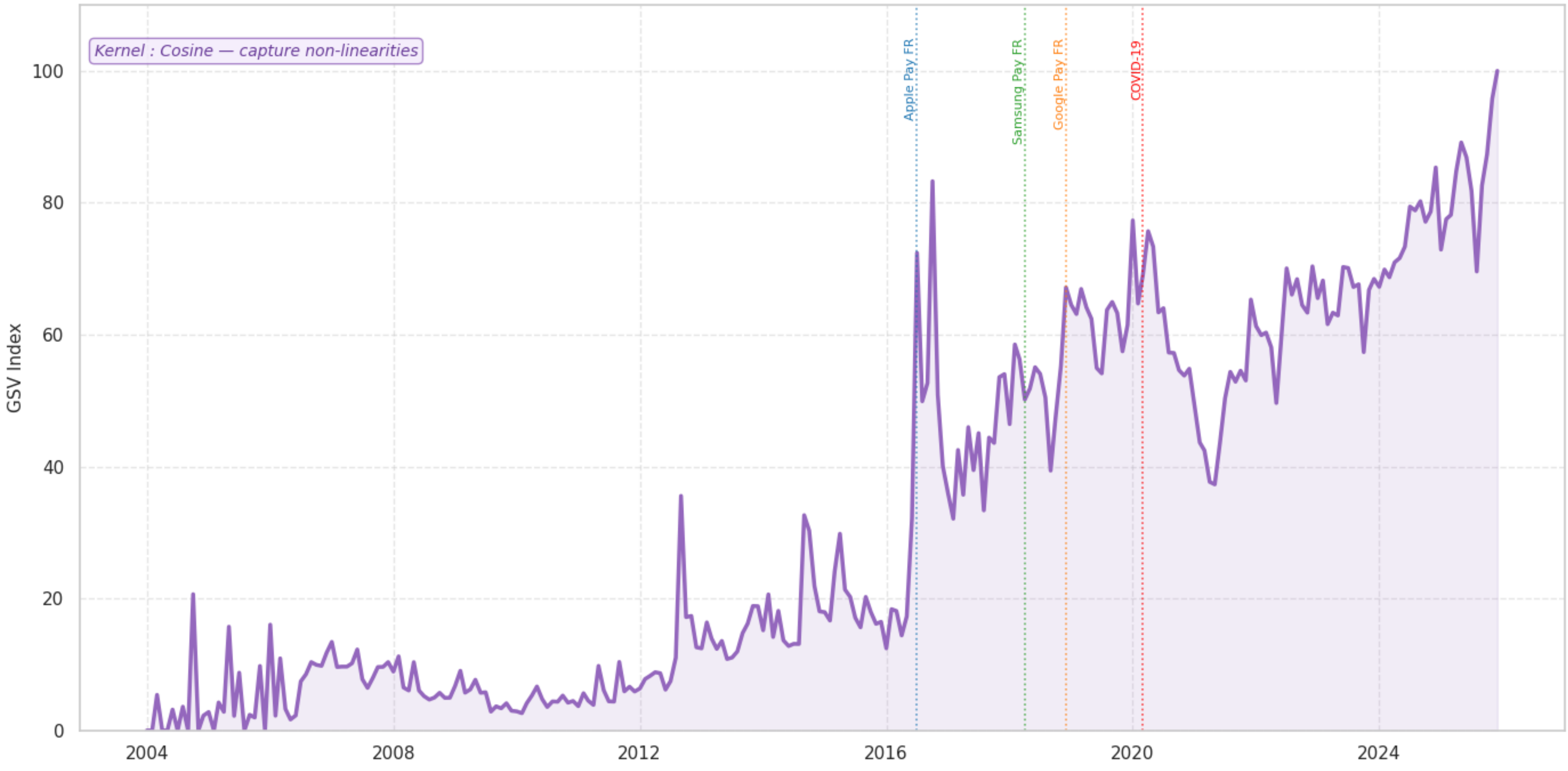
LSTM autoencoder

- LSTM autoencoder is a specialized Recurrent Neural Network (RNN) designed to compress complex, multivariate time-series data.
- Analyzes historical sequences to capture temporal dynamics and cross-trend correlations.
- Forces multi-dimensional inputs into a single, highly condensed vector.
- Unifies raw search sequences into core underlying consumer payment behaviors over time.

Kernel PCA index

- Maps the original data into a higher-dimensional feature space using a kernel function and then performs PCA in this new space

Chart 7: Kernel PCA Index of digitization

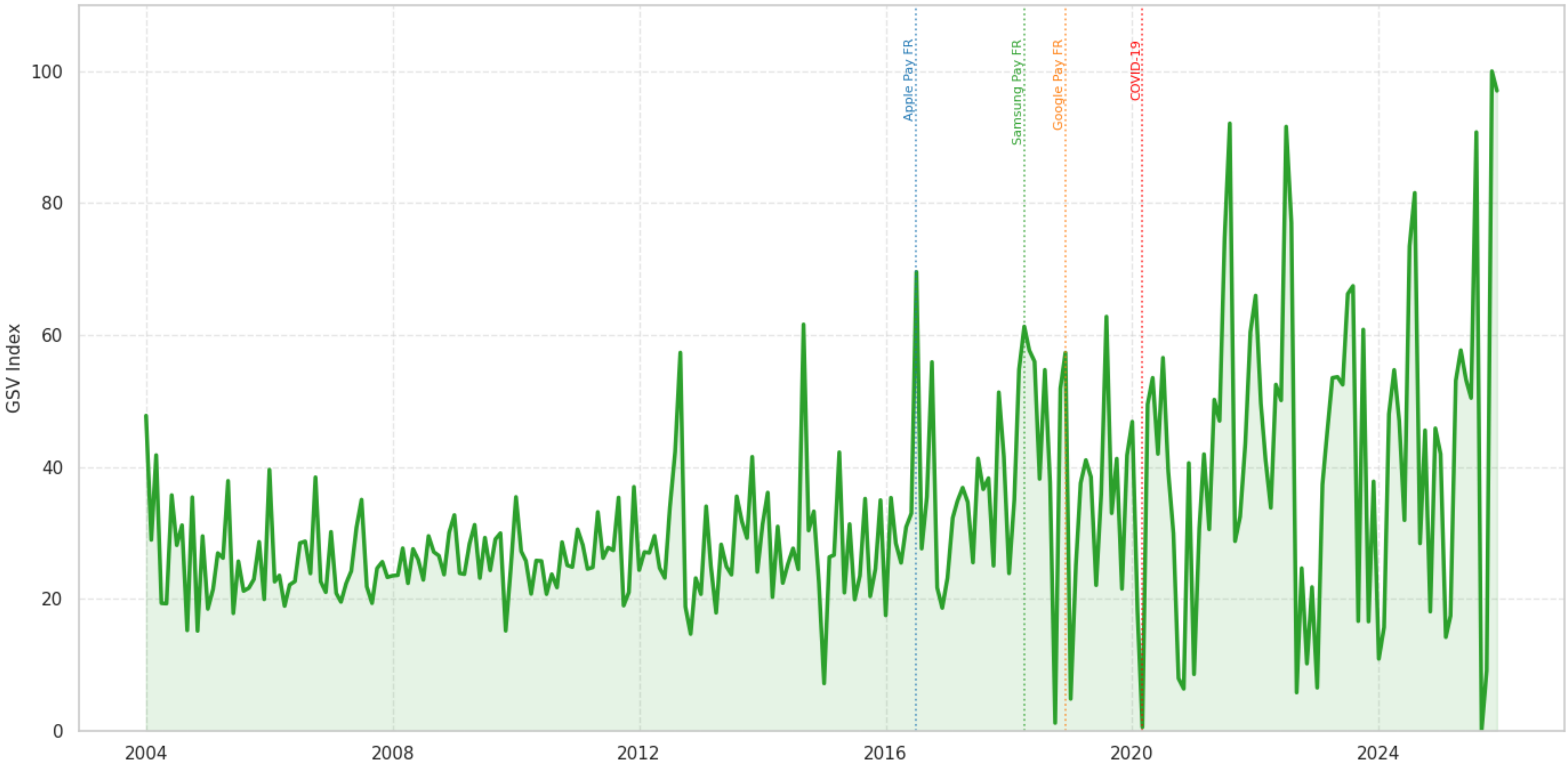


Source: author's calculations

Temporal dynamics: the DFM model

- The DFM indicator registers sharper and earlier responses to the launch of mobile payment platforms and to the COVID-19 shock, reflecting the Kalman filter's ability to update the latent factor estimate in real time as new search patterns emerge

Chart 8: DFM indicator

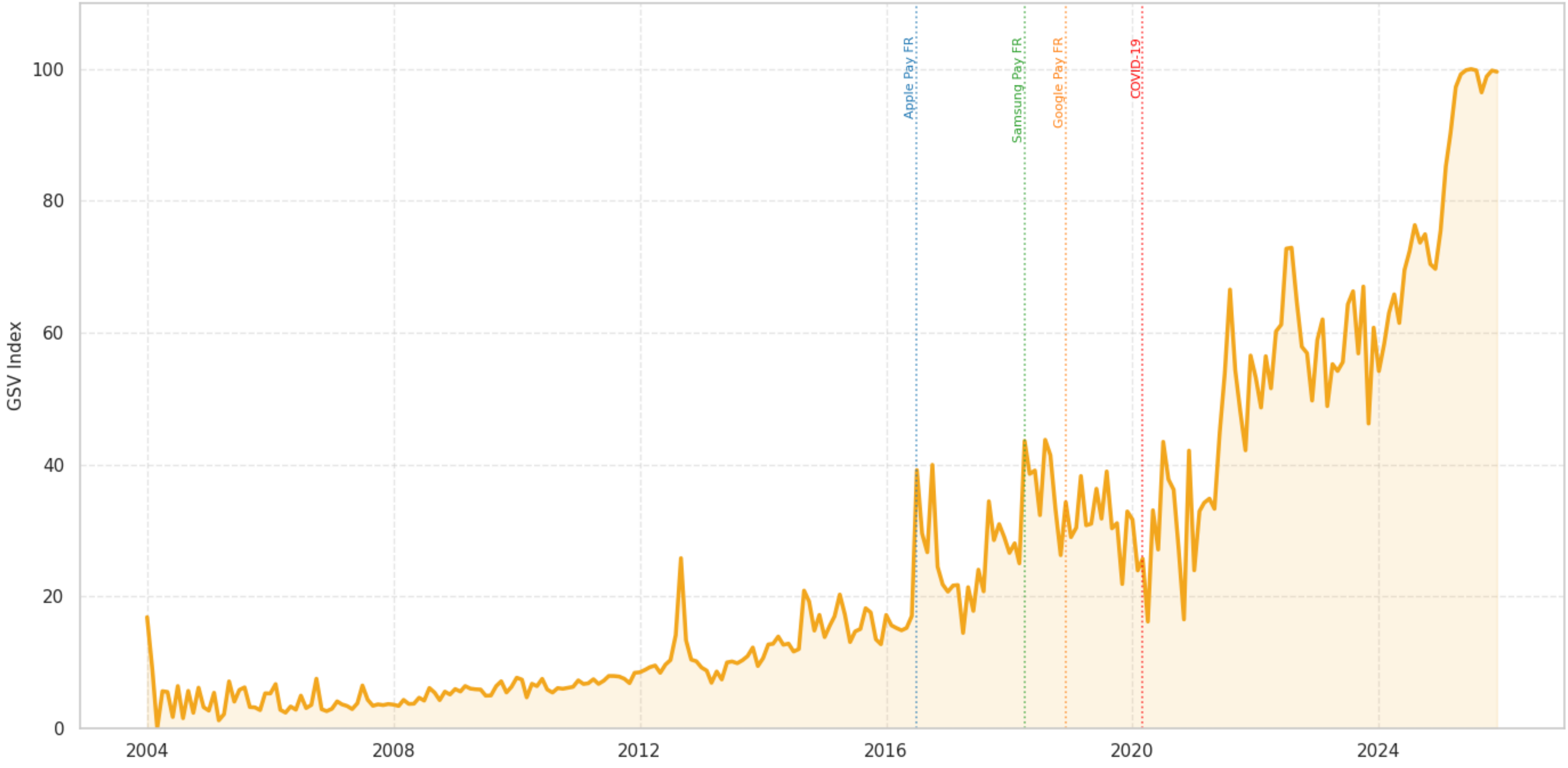


Sources: author's calculations

Sequential dependence: LSTM autoencoder

- LSTM autoencoder compresses each month's search vector into a latent state, explicitly modeling memory and ordering effects that a static factor model ignores.

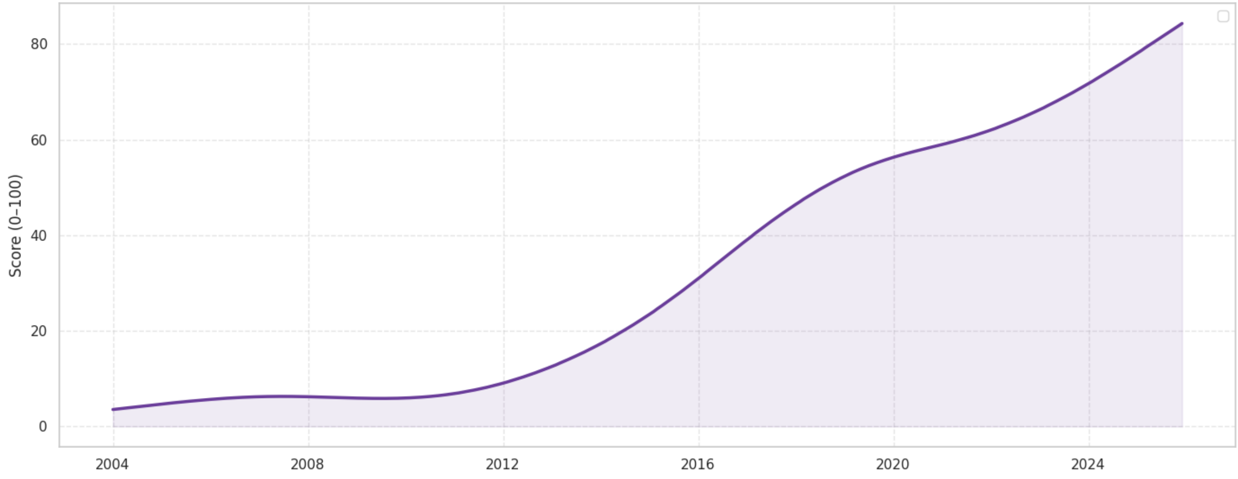
Chart 9: LSTM Autoencoder indicator



Sources: author's calculations

Chart 10: Trend Component (τ_t) from Hodrick-Prescott Filter on build Payment Digitization Indicator

a) Kernel PCA

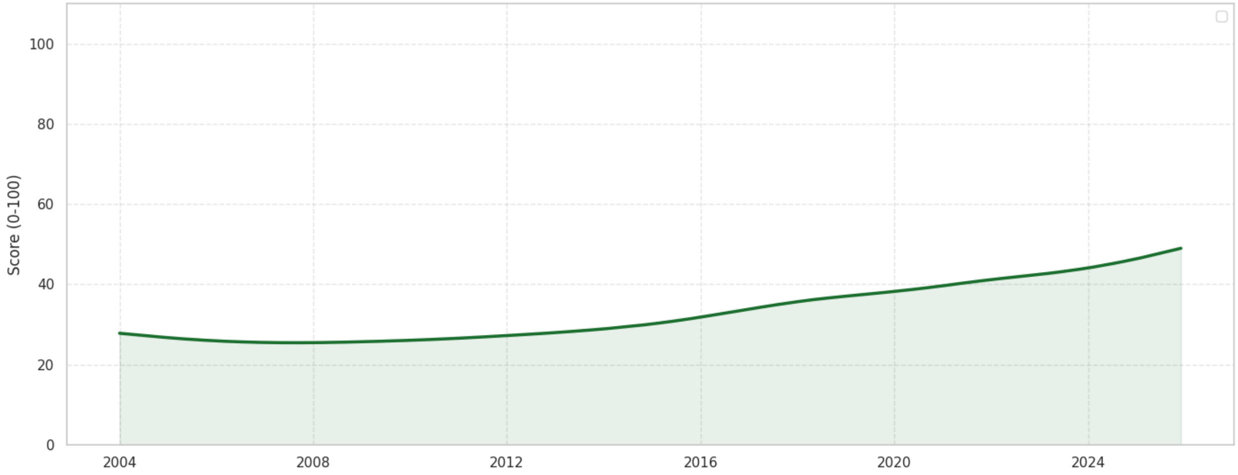


Use of the Hodrick-Prescott (HP), mathematical smoothing technique, to identify two unobserved components:

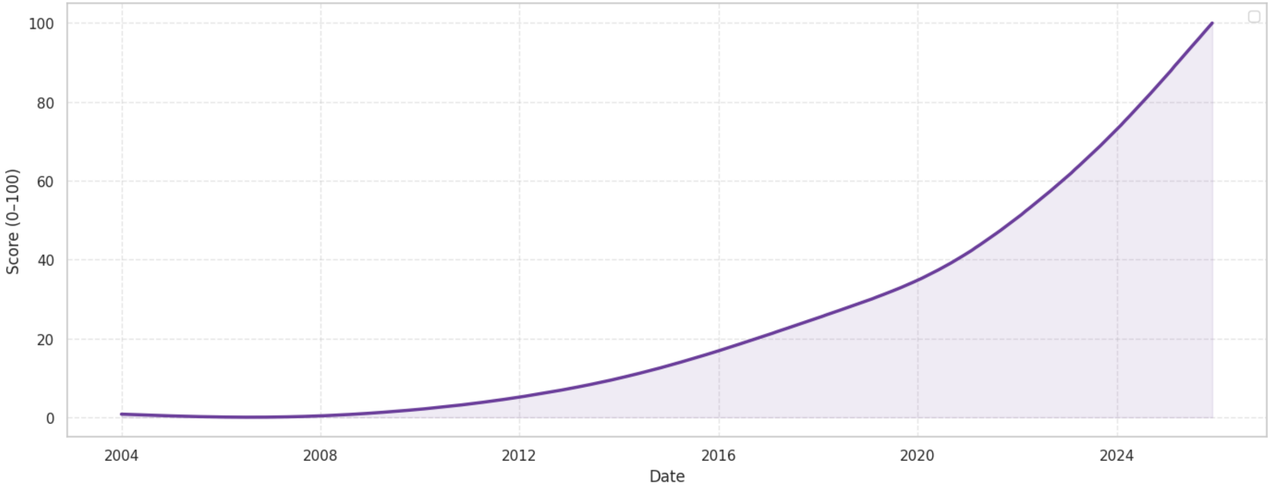
- Trend component (τ_t): The smooth, long-term growth path.
- Cyclical component (c_t): Short-term fluctuations around the trend (business cycle deviations).

→ The results show a more pronounced upward trend in the Kernel PCA and LSTM Autoencoder indicators than in the DFM, which is flatter.

b) DFM



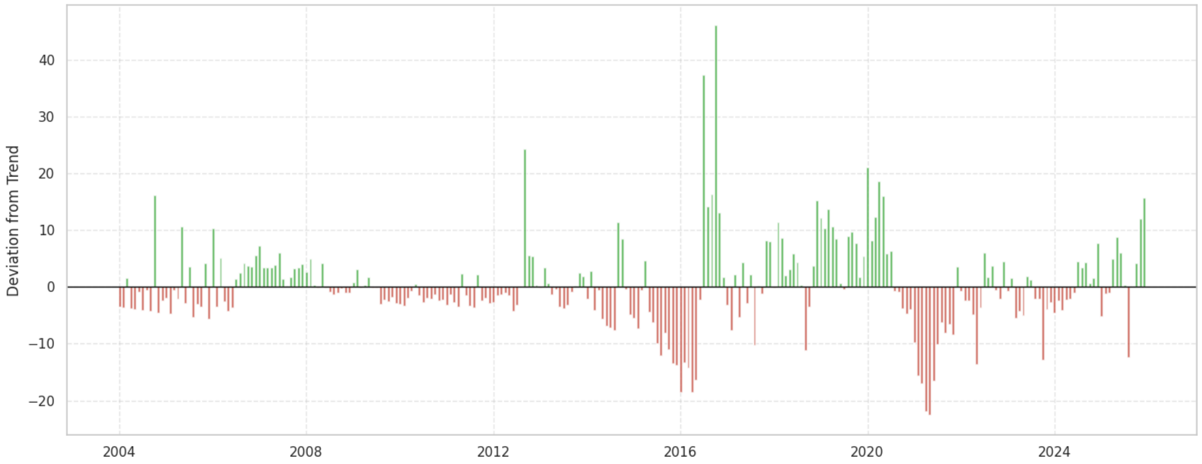
c) LSTM Autoencoder



Sources: author's calculations

Chart 11: Cyclical Component (c_t) from Hodrick-Prescott Filter on build Payment Digitization Indicator

a) Kernel PCA

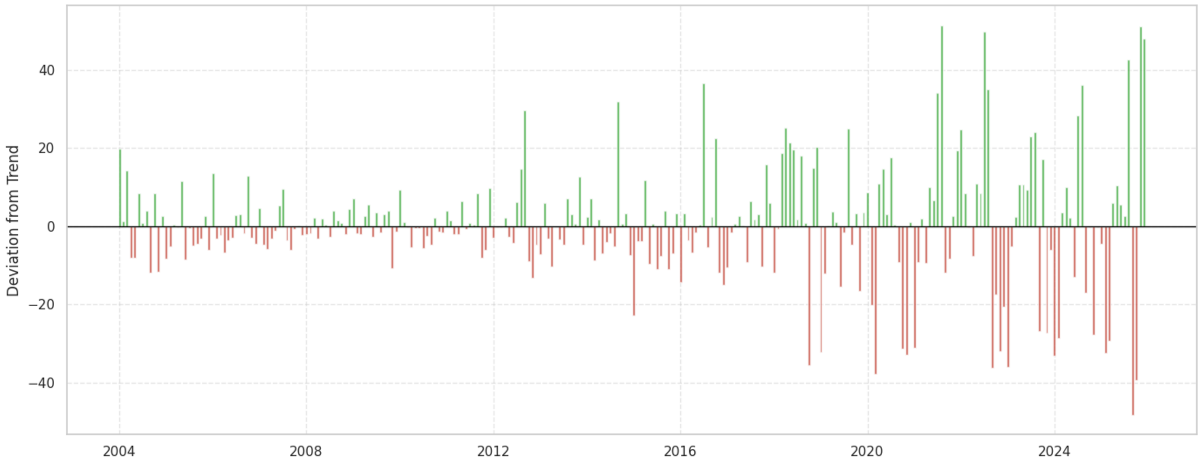


Use of the Hodrick-Prescott (HP), mathematical smoothing technique, to identify two unobserved components:

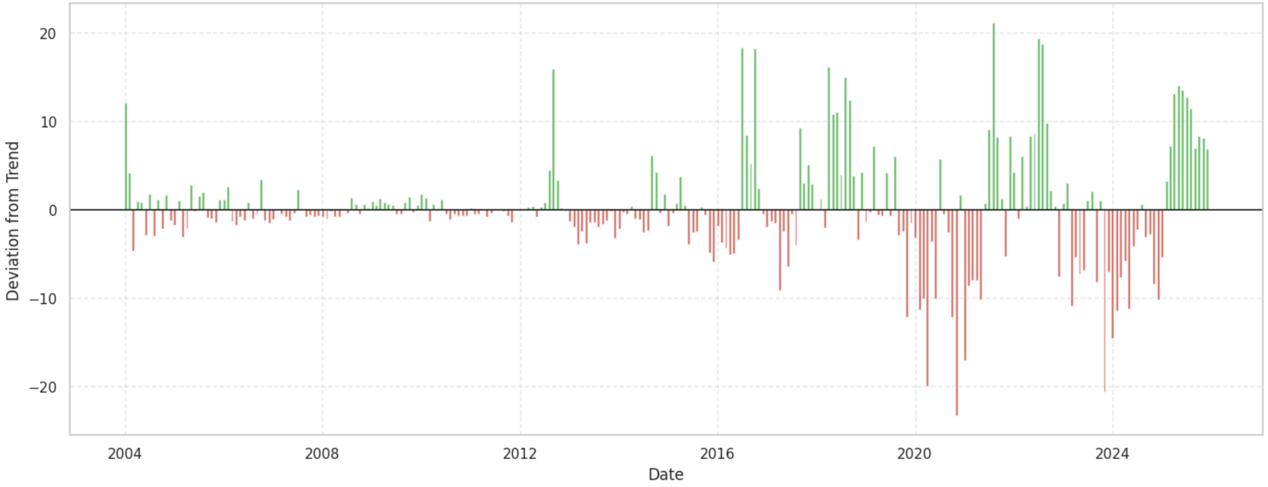
- Trend component $(\boldsymbol{\tau_t})$: The smooth, long-term growth path.
- Cyclical component (c_t): Short-term fluctuations around the trend (business cycle deviations).

→ The DFM cycles are more erratic than the Kernel PCA and LSTM indicators.

b) DFM



c) LSTM Autoencoder



Sources: author's calculations

Conclusion and next steps

- Evaluating Three Aggregation Methodologies for Google Search Volume Data.
- The Granger causality test shows bidirectional causality between DFM and LSTM.

Next steps :

- Assess consistency of used methodology and hypothesis
- Test coherence consistency with macroeconomic indicators: concordance tests, correlations tests, granger causality tests between the GSV-DPI and macroeconomic; monetary/financial aggregates, dynamic correlation analysis over the full sample period.

THANK YOU FOR YOUR ATTENTION

ANY QUESTIONS ?

Louisalexandre.bayol@banque-france.fr

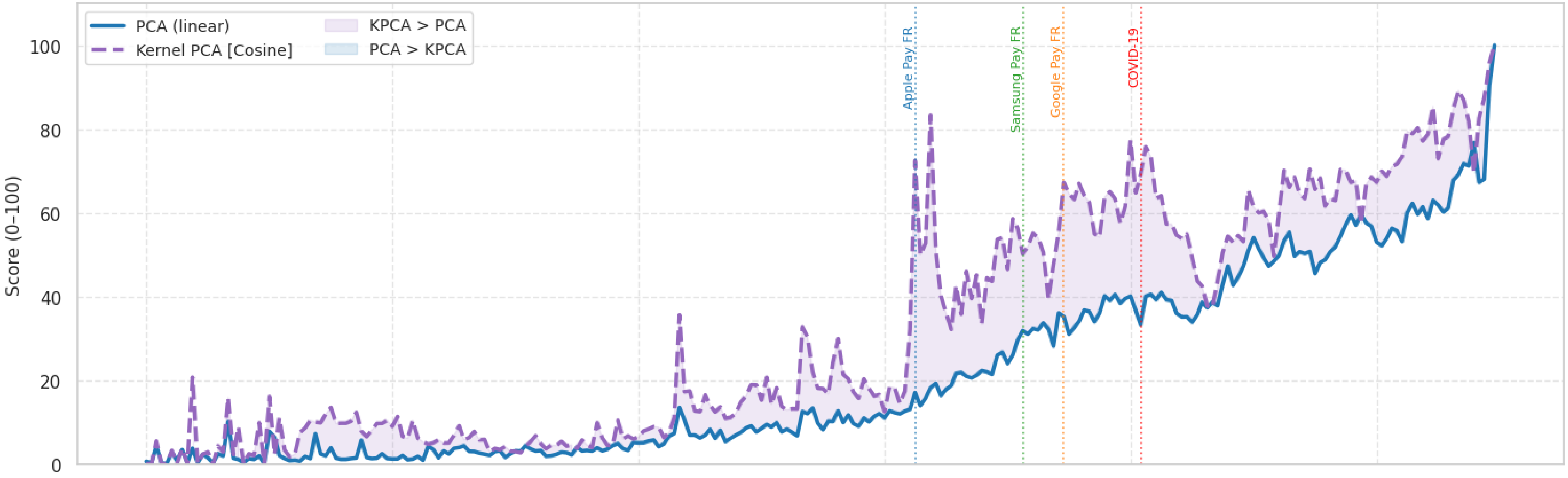
Ageraut@inseec.com

- Avouyi-Dovi, Sanvi, Bayol, Louis-Alexandre, Devigne, Lucas, Geraut, Ambre and Pfister, Christian (Sept., 2026). "Digitization of payments in France: Lessons from large language models". In: *Quarterly Review of Economics and Finance 109 (2026) 102176.*
- Askitas, Nikos and Klaus F. Zimmermann (June 1, 2009). "Google Econometrics and Unemployment Forecasting". Rochester, NY". (Visited on 06/25/2026).
- Bikourne, Mariem et al. (Jan. 12, 2025). "Statistical methods for inflation forecasting in Morocco: Insights from Google trends data". In: *Statistics, Optimization & Information Computing 13.4, pp. 1636–1653.* (Visited on 06/25/2026).
- Choi, Hyunyoung and Hal Varian (2012). "Predicting the Present with Google Trends". In.
- D'Amuri, Francesco and Juri Marcucci (2012). "The Predictive Power of Google Searches in Forecasting Unemployment". In: *SSRN Electronic Journal. (Visited on 06/25/2026).*
- Da, Zhi, Joseph Engelberg, and Pengjie Gao (Oct. 2011). "In Search of Attention". In: *The Journal of Finance 66.5, pp. 1461–1499.* issn: 0022-1082, 1540-6261. (Visited on 06/25/2026).
- Diebold, Céline (June 2025). "Using Google search data to examine factory automation and its effect on employment". In: *Economic Analysis and Policy 86, pp. 1301–1328.* (Visited on 07/21/2026).

- Eichenauer, Vera Z. et al. (2021). “Obtaining consistent time series from Google Trends”. In: *Economic Inquiry 60(2):694-705.*
- Götz, Thomas B and Thomas A Knetsch (2019). “Google data in bridge equation models for German GDP”. In: *the International Journal of Forecasting (Vol. 35, Issue 1, pp 45–66).*
- Schmidt, Torsten and Simeon Vosen (2010). “A Monthly Consumption Indicator for Germany Based on Internet Search Query Data”. In: *Ruhr Economic Paper No. 208.*
- Stock, James H. and Mark W. Watson (2002). “Forecasting Using Principal Components from a Large Number of Predictors”. In: *Journal of the American Statistical Association 97.460, pp. 1167–1179.* (Visited on 06/25/2026).
- Tenorio, Juan et al. (Mar. 27, 2025). “An Artificial Trend Index for Private Consumption Using Google Trends”. In: *Apuntes 99, 5-45. (Visited on 06/25/2026).*
- Varian, Hal R. and Hyunyoung Choi (Apr. 2, 2009). “Predicting the Present with Google Trends”. In: *Rochester, NY.* (Visited on 06/17/2026).

Comparison of Kernel PCA and PCA

Chart A.1 : Non-linearity captured by Kernel PCA vs PCA



Sources: author's calculations