The Impact of Artificial Intelligence on the Consumer Confidence Index: Results from a Heterogeneous Panel Data Analysis

Nasimi Nuriyev[1,2,]
E-mail: nesimi.nuriyev@ogr.iu.edu.tr

Dundar Murat Demiroz[1]
E-mail: dundar.demiroz@istanbul.edu.tr

Hasraddin Guliyev[2,]
E-mail: hasradding@unec.edu.az

Abstract

Homoeconomicus, the fundamental principle of economics, posits that decisions are based on past experiences and anticipated future outcomes. A plethora of factors can influence consumer behavior. Micro-level personal experiences, macroeconomic indicators, and societal performance all influence consumers' perceived trust, a crucial determinant of their expectations Twitter is one of the most extensive online platforms for information exchange. Since its inception, researchers have explored the potential of using Twitter data in various decision-making scenarios, which has yielded promising results. Our research, based on a comprehensive literature review, aims to highlight the factors that predict and influence consumer interactions with AI posts about artificial intelligence. Using a heterogeneous panel data model, we analyzed the Consumer Confidence Index (CCI), tweets about artificial intelligence on Twitter (AItwit), Consumer Price Index (CPI), Long Term Interest Rates (LTIR), non-OECD economies (4 countries), and OECD economies (24 countries). This model spanned a period of 27 months (January 2021–March 2023) and covered a total of 28 countries. We then employed the Extended Mean Group (AMG) joint process, a heterogeneous panel estimator with autocorrelation, developed by Eberhardt and Teal (2010) and Eberhardt and Bond (2009). We calculated the difference between the dependent variable and the classifier-lasso model, as suggested by Su, Shi, and Phillips (2016). We found a statistically significant econometric relationship between all variables.

Keywords: Artificial intelligence tweets on Twitter, Consumer confidence index, consumer price index, Long-term interest rates, heterogeneous panel, Extended Mean Group (AMG) and Classifier-lasso

JEL Classification: D12, D84, E20, E30, E40, C23

1. Introduction

Twitter has become a crucial tool for disseminating information and has significantly influenced consumer behavior and economic patterns in the digital era. Twitter's extensive use as a social media platform has created a conducive environment for the easy acquisition and dissemination of diverse viewpoints and knowledge. For instance, studies have observed substantial volatility in stock prices due to the social media posts made by CEOs such as Elon Musk (Ante L., 2023; Molla R., 2021). With the growing significance of social media in the economy, consumers are becoming

more aware of the need for monitoring and analyzing social media data. This enables them to make educated decisions and foresee economic trends. While the proliferation of new technology presents new prospects for companies, it also gives rise to adverse repercussions (DEMİRÖZ, D., 2019). Technology capacity techniques assess developing nations' ability to improve their technological skills, taking into account production conditions (Tahsin, E., et al., 2023). The integration of AI into people's daily lives can assume many roles in the connection between consumers and AI, spanning from basic transactional exchanges to more intricate empathetic partnerships.

The main goal of the current study is to analyse the impact of AI-related tweets on the consumer trust index in order to determine the factors that influence the connections between consumers and artificial intelligence (AI) (Olan F. et al., 2021). Integrating AI into daily life may lead to many types of interactions between consumers and AI, spanning from simple transactional ties to more intricate empathetic bonds. The progress in AI technology and its applications has significantly altered consumer behaviour, as people engage with these apps on many platforms and touchpoints. Understanding the impact of these encounters on customer behaviour, including aspects like personality, attitude, engagement, decision-making, and trust, is of utmost importance (Lain V. et al., 2024).

In addition, the integration of artificial intelligence (AI) into service and industrial robots is changing customer behaviour and satisfaction, highlighting the importance of advanced technologies in economic decision-making (Corina P. et al., 2021). Moreover, the automation bias illustrates that customers are increasingly dependent on AI advice rather than human knowledge, even in the face of unsatisfactory outcomes. This might potentially influence how consumers perceive economic trends and make financial decisions.

The advent of productive AI has elicited positive sentiments, particularly among individuals who have had interactions with AI, and there has been a significant upsurge in interest following the launch of ChatGPT (Miyazaki, K., et al., 2024). The advancement of artificial intelligence has engendered apprehensions about potential manipulation and dependence on intelligent technologies, which can impact consumer decision-making and cognitive capabilities (Linwan Wu et al., 2021).

Consumers represent a significant economic decision-making unit in macroeconomics. Consumer behavior has a considerable impact on decision-makers and economic forecasts. The factors influencing consumer behaviour are numerous. Consumers' expectations and the trust between economic agents shape their behaviour. Consumers' expectations for the future may result in different economic outcomes. For instance, optimistic expectations may prompt consumers to spend more and borrow more, while pessimistic expectations may lead them to reduce their expenditures and reassess their financial situation. The consumer confidence iWe use the consumer confidence index to interpret consumption and other macroeconomic indicators. . (2001) frequently employ it to gauge investor sentiment.

According to a study by Millet B. et al. (2024), individuals who perceive Twitter as useful for information are more likely to trust social media news and engage with it by liking or sharing it.

This indicates that people's perceptions of Twitter's usefulness may significantly influence how they interact with online news. The insights from this research can be beneficial for future studies on the dissemination and comprehension of news on social media platforms like Twitter.
Consumer trust is considered a subject within the domain of psychological economics. Consumer confidence is dependent on both purchasing power and willingness to buy household goods. The theoretical framework that seeks to comprehend alterations in behavior is dynamic theory, which should track changes in consumer tastes, preferences, and attitudes (Katona G. 1968). Consumer behavior holds significant weight in the deliberations of economic decision-makers and serves as a critical source of information for economic forecasters regarding future projections.
The main thesis posits the significance of the Twitter discussion on economic AI in shaping consumer behavior and expectations, and the need to balance optimistic sentiments with concerns about AI's effects on the economy and society. This paper aims to contribute to the ongoing debate with a more in-depth perspective and comprehensive empirical analysis in the context of both OECD member and non-member countries.
The paper organizes the subsequent sections as follows: Section 2 provides a review of relevant studies. Section 3 outlines the methodologies used in the analysis. Section 4 presents the dataset, variables, and empirical results. Lastly, Section 5 offers discussion and policy implications.

2. Literature Review

Research on AI in online social networks has yielded significant findings. It is vital to recognize the role that both consumers and service providers play in the dissemination of knowledge. However, it is equally important to understand how the subnetwork of consumers and service providers within larger social networks, such as Twitter, contributes uniquely to this process. Twitter also serves as a platform for expressing opinions about competitors in any industry. The views of competitors can sway the sentiment and decision-making of both consumers and service providers.
Bodaghi et al. (2020) conducted a review of 28 articles over a five-year period to examine the self-mechanism, findings, and methods, including problem formulation, solutions, Twitter features, and knowledge transformation. The study aimed to evaluate the impact of Twitter's citation promotion policy during the 2020 US presidential election using a large dataset of 304 million tweets from 86,334 users. The researchers used SARIMAX models, linear regression, and text and sentiment analyses to analyse how the policy affected retweeting and quoting behaviour at different intervals aligned with the timeline. During the policy's introduction, the results showed a noticeable decline in retweeting activity and a modest increase in quoting behavior, followed by a rapid resurgence in retweeting and a decline in quoting after the policy period. These findings provide valuable insights for social media policymakers to improve existing policies and implement more effective interventions based on observed user behaviour and preferences.
Cao et al. (2023) used Twitter's semantic space to predict financial anomalies by employing almost two million entities extracted from news texts disseminated through popular news channels over a

12-year period. The study concentrates on the evolution of assets in semantic graphs to accurately predict future stock market anomalies. This study employs centrality measures within the daily semantic graphs of these entities to identify highly correlated entities, which are then utilized to construct a logistic regression model. The resulting model demonstrates remarkable accuracy in predicting the emergence and direction of future anomalies in stock prices, whether indicating an upward trend for prices to rise or a downward trend for prices to fall in the future. The research illustrates the interdisciplinary nature of the field by integrating natural language processing, complex networks, and artificial intelligence techniques to formulate innovative forecasting models that offer substantial advancements in predicting financial anomalies. The study's outcomes hold great potential for future research endeavours due to the proficient application of sophisticated analytics to forecast stock market anomalies, thereby contributing to the development of forecasting models in finance.

As per Bianchi et al. (2024), individual politicians have the ability to impact asset prices by utilising social media. The sentiments conveyed in tweets by members of Congress can directly influence the stock prices of targeted companies and industries. Furthermore, when tweets contain information about upcoming legislative activity, the effects on stock prices intensify. The findings illustrate that "supportive tweets lead to an increase in the stock prices of the targeted firm and the relevant industry, while critical tweets result in a decrease in the same within minutes of the tweet's posting."

Bianchi et al. (2023) make a significant contribution by identifying threats to central bank independence from high-frequency financial data and social media messages, particularly from the president's account. This methodology sets the study apart from prior research that focuses on constructing independence indices across countries. The study emphasizes the importance of monitoring social media platforms, such as Twitter, for early signs of attempts to undermine central bank autonomy and provides a valuable tool to enhance the resilience of monetary policy frameworks. The findings of this research suggest a shift towards a more holistic approach to assessing and safeguarding central bank independence, highlighting that central banks should consider not only traditional economic variables but also sentiments expressed on social media.

Researchers have conducted numerous investigations to determine the impact of consumer trust. According to Mehra and Martin's (2003) findings, consumer trust holds considerable importance in regression models designed to predict consumer expenditure. Garrett, Hernández-Murillo, and Owyang (2005) analysed regional data and provided evidence that consumer confidence can serve as a dependable predictor of retail expenditure in various states of the United States. Vuchelen (2004) looked at the relationship between consumer sentiment, anticipated income, and the degree of uncertainty associated with that income.

Shayaa et al. (2018) investigated the connection between the Consumer Confidence Index (CCI) and social media sentiment analysis in comprehending consumer purchasing behavior over a two-year period from January 2015 to December 2016. The results indicated that social media sentiment analysis can complement conventional methods of analysing consumer confidence by providing a quicker and potentially more accurate means of measuring consumer behaviour by

utilising big data from a broader population via platforms such as Twitter. Grasping the relationship between consumer confidence and social media sentiment can aid businesses and policymakers in making better-informed decisions about marketing strategies and economic policies. By examining social media data alongside traditional economic indicators like the Consumer Confidence Index (CCI), companies can gain more timely and precise insights into consumer preferences and sentiments and adjust their products and services accordingly.

In Japan, where Golinelli and Parigi (2004) conducted their research, the government debt-to-GDP ratio and unemployment rate significantly influence consumer confidence. It is believed that a high unemployment rate negatively affects mental health through job insecurity (Erlinghagen, 2008; De Witte et al., 2016). Acemoglu and Scott (1994) and Carroll, Fuhrer, and Wilcox (1994) discovered that past consumer sentiment has a significant influence on current changes in household spending in both the US and the UK. Kuhnen and Knutson (2011) found in their study on economic risk assessment and decision-making processes that participants exposed to negative photographs preferred risk-free assets over risky ones. To assess economic well-being, economic forecasters use consumers' personal financial expectations and their views on the overall state of the macroeconomy.

3. Methodology

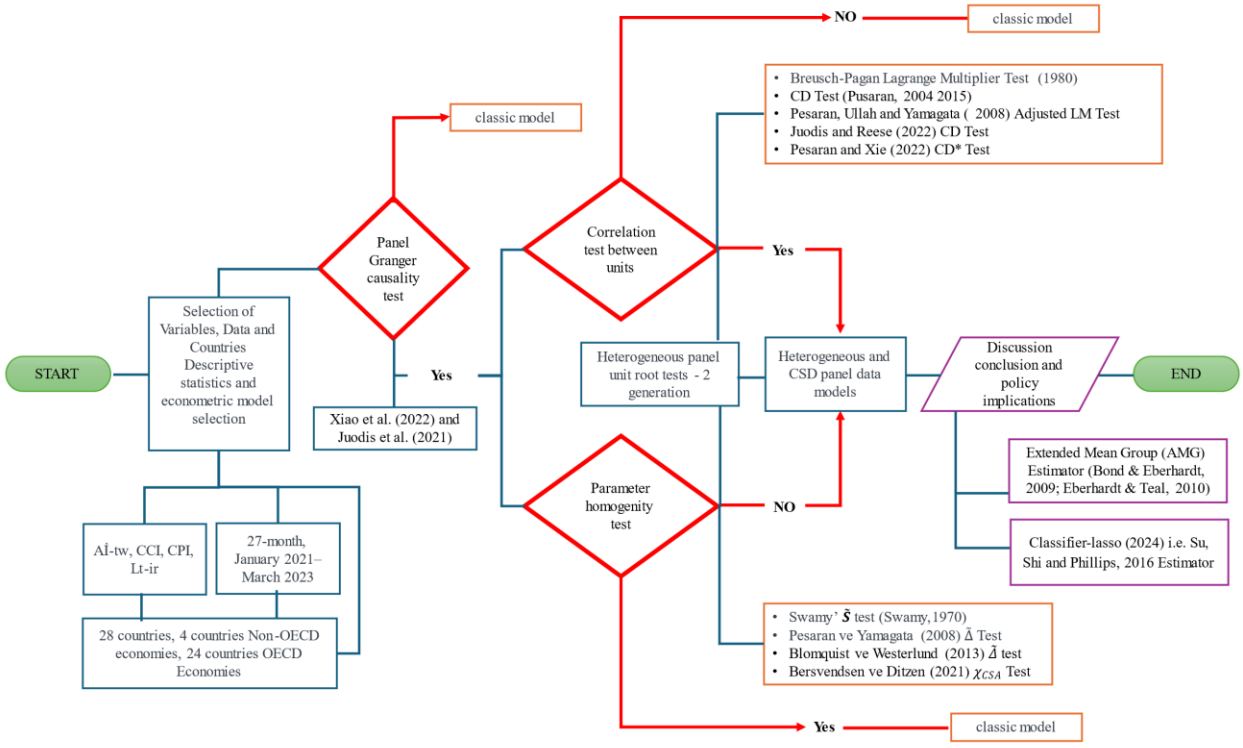
Panel data analysis is a method of estimating economic relations using panel data models constructed with horizontal cross-sectional data that includes a time dimension (Yerdelen Tatoğlu, 2020:37). Unlike a normal horizontal-section or time-series regression, a panel data regression has both unit and time dimensions, which are represented by a pair of subscripts over the variables (Baltagi, 2021:15). The standard panel data regression model is typically depicted in the following manner:

$$y_{it} = \alpha + \sum_{j=1}^{k} \beta_j X_{jit} + u_{it} \qquad (1)$$

Here i denotes units such as households, individuals, firms and countries and t denotes time. Therefore, sub-index i indicates the horizontal cross-section dimension and t indicates the time series dimension in the data. In the context of equation (1), it is notable that both the fixed parameter and the slope parameters are homogeneous and do not vary across units or time. Models that exhibit this property are known as pooled regression models. The Pooled Least Squares (HEKK) estimator is commonly utilized for the estimation of these models (Gujarati et al., 2012: 640645). To address the challenges posed by "long" panels with high time dimensionality, researchers developed heterogeneous panel data models. These models have become increasingly popular in empirical research across many fields, including development economics, regional science, and political science (Guliyev.H.,2022). These datasets aggregate a large number of time series from different countries or firms. However, in many applications, the response of a dependent variable to changes in explanatory variables may differ across all units or over time. Accounting for heterogeneity in the interaction between variables in panel data analysis can help

advance econometric research and better understand real economic phenomena. Heterogeneous panel data models, in contrast to classical panels, place emphasis not only on the asymptotics of N but also on the asymptotic convergence of both N and T towards infinity, either sequentially or jointly (Phillips and Moon, 1999). Often, we observe that the individual coefficients of the units can vary across units, resulting in heterogeneous specifications. Therefore, such panels challenge the assumption of parameter homogeneity. In these models, it is theoretically possible to classify parameter heterogeneity as one-way or two-way heterogeneous panel data models. The flowchart provided below succinctly captures the analysis process involving the heterogeneous panel model.

Figure 1: Flowchart summarising the analysis of the heterogeneous panel model



3.1. Choice between Homogeneous and Heterogeneous Panel Data Models

When choosing between homogeneous and heterogeneous panel data models, it is critical to first determine whether the panel data structure is appropriate for a homogeneous or heterogeneous structure. Upon detecting heterogeneity in the panel data model through the slopes, the subsequent step involves ascertaining the presence of an inter-unit correlation among the panel's cross-sections. If the panel data model detects no inter-unit correlation, the RCM or MG estimators are the appropriate estimators. On the other hand, if inter-unit correlation is detected in the model, the SUR, CCE, and AMG estimators are more appropriate (Yerdelen Tatoğlu, 2020: 101). It is important to first check the validity of homogeneity based on the slopes and then see if there is a correlation between the units in order to make the right choice of estimator in heterogeneous panel data models.

Panel data models use homogeneity tests to evaluate the homogeneity of the slope parameters. In other words, these tests aim to determine whether the coefficients calculated for each panel are equal to one another. The main hypothesis in homogeneity tests is that the coefficients calculated for all units are equal to each other, whereas the alternative hypothesis is that the coefficients calculated for at least one unit are different from the others (Pesaran, 2015: 734).

H_0: $\beta_i = \beta$ (all i for all units) Parameters for all units are homogeneous

H_1: $\beta_i \neq \beta$ (any i for) Parameters are not homogeneous (heterogeneous)

Swamy's (1970) $\tilde{S}$ Test: Swamy (1970) proposed a test known as the statistic to determine the homogeneity of the slope coefficients. In the $\tilde{S}$ test, the unit dimension N is required to be smaller than the time dimension T, and it is a test resistant to heteroskedasticity across units. To calculate Swamy's (1970) $\tilde{s}$ statistic, follow these steps:

$$\tilde{S} = \sum_{i=1}^{N} \left(\hat{\beta}_i - \hat{\beta}_{WFE}\right) \frac{X_i' M_\tau X_i}{\hat{\sigma}_i} \left(\hat{\beta}_i - \hat{\beta}_{WFE}\right) \quad (2)$$

In the equation, $\hat{\beta}_i$ is the pooled OLS estimator, $\hat{\beta}_{WFE}$ is the pooled estimator of the weighted fixed effect model, M_τ is the unit matrix, and $\hat{\sigma}_i$ is the estimator of σ_i. The test statistic has a Chi-squared distribution with K×(N-1) degrees of freedom. If the calculated $\tilde{s}$ statistic is greater than the critical value, the null hypothesis is rejected, and the slope parameters are heterogeneous for at least one unit (Yerdelen Tatoğlu, 2020: 97).

Pesaran and Yamagata (2008) $\tilde{\Delta}$ Test: A standardised version of the Swamy (1970) $\hat{S}$ statistic, which is valid for panels with large unit (N) and time (T) sizes. The first process of this test, called $\tilde{\Delta}$, starts with the calculation of the $\hat{S}$ statistic, and the $\hat{S}$ statistic is adjusted in the next stage (Pesaran, 2015: 738–739).

$$\Delta = \sqrt{N}\left(\frac{N^{-1}\tilde{S}_i - k}{\sqrt{2k}}\right) \quad (3)$$

This test is valid as $(N,T) \to \infty$ provided $\sqrt{N}/T^2 \to 0$ and assumes that the error terms are asymptotically normally distributed $\tilde{\Delta} \sim N$ and uncorrelated across units and time (ε_{it} ile $\varepsilon_{js}, i \neq j$ ve $t \neq s$). The test allows for unit-level heteroskedasticity and is also valid for first-order autoregressive models (dynamic panel data models), but for validity, it is expected that the unit time dimension is close in size ($N/T \to k$) and rapidly converges to infinity at the same time $N,T \to \infty$. In practice, especially when faced with panels of small units (N) size, it is unlikely that N and T will converge to infinity at the same speed, and thus the test loses power. Therefore, Pesaran and Yamagata (2008) proposed a corrected version of the test, the $\tilde{\Delta}_{adj}$ test:

$$\tilde{\Delta}_{adj} = \sqrt{N}\left(\frac{N^{-1}\tilde{S} - k}{\sqrt{\mathrm{var}\left(\tilde{z}_{i,T}\right)}}\right) \quad (4)$$

Blomquist ve Westerlund (2013) $\tilde{\Delta}$ test: $\tilde{\Delta}_{\text{HAC}}$ test by modifying it to be robust to autocorrelation and heteroskedasticity.

$$\tilde{\Delta}_{\text{HAC}} = \sqrt{N}\left(\frac{N^{-1}S_{HAC} - k}{\sqrt{2k}}\right) \textbf{Here it is } S_{HAC}$$

$$= \sum_{i=1}^{T} T\left(\hat{\beta}_i - \hat{\beta}\right)'\left(\hat{Q}_{i,T}\hat{V}_{i,T}^{-1}\hat{Q}_{i,T}\right)\left(\hat{\beta}_i - \hat{\beta}\right) \quad (5)$$

$\hat{\beta}$ is the ECM estimator for each unit and $\hat{\beta}$ is calculated by equation (4):

$$\hat{\beta} = \sum_{i=1}^{N}\left(T\hat{Q}_{i,T}\hat{V}_{i,T}^{-1}\hat{Q}_{i,T}\right)^{-1}\sum_{i=1}^{N}\hat{Q}_{i,T}\hat{V}_{i,T}^{-1}X_i'M_\tau y_i \quad (6)$$

In equation (5) X_i' is a $T \times k$ dimensional matrix of independent variables, $M_\tau = I_T - \tau_t(\tau_t{}'\tau_t)^{-1}\tau_t'$ being equal to $\tau_t, T \times 1$ dimensional unit matrix, $\hat{V}_{i,T_i}$ is a robust estimator based on kernel density function.

Bersvendsen and Ditzen (2021) $\tilde{\Delta}_{CSA}$ Test: Test developed by Pesaran and Yamagata (2008) $\tilde{\Delta}$ test and the test modified by Blomquist and Westerlund (2013) to be robust to autocorrelation and heteroskedasticity $\tilde{\Delta}_{HAC}$ test does not take into account the correlation between units. Bersvendsen and Ditzen (2021), $\tilde{\Delta}$ investigated the power and size properties of the test against inter-unit correlation. The CSD Resilient Delta test developed by Bersvendsen and Ditzen (2021) is a heterogeneity test of the slope parameter extended to the DCCE estimator. Available $\tilde{\Delta}$ partially disabled the DCCE estimator to eliminate the strong inter-unit correlation in the model in order to make the test robust: $\nabla_t = \left(\tilde{v}_t, \ldots, \tilde{v}_{t-p}\right)(6)$. In equation (6) $\overline{v_t}$ It is a vector consisting of the horizontal cross-sectional mean and lags of the dependent and independent variables. Using equation (6), the following process is applied and the variables in the model are deflated:

$$M_{\overline{v_t}} = I_T - \nabla_t(\bar{V}_t'V_t)^{-1}V_t' \quad (7)$$

$$\tilde{Y}_{it} = Y_{it}M_{\overline{v_t}}, \tilde{Y}_{i,t-1} = Y_{i,t-1}M_{\overline{V_t}} \text{ ve } \tilde{X}_{it} = X_{it}M_{\overline{V_t}}$$

Using the variables obtained from equation (7) $\tilde{\Delta}$ and $\tilde{\Delta}_{HAC}$ test processes are applied and $\tilde{\Delta}_{CSA}$ test statistic is calculated (Bersvendsen and Ditzen, 2021: 51-80). As a result of their Monte Carlo simulations, Bersvendsen and Ditzen (2021) use the test statistic developed by Blomquist and Westerlund (2013) against heteroskedasticity and autocorrelation. $\tilde{\Delta}_{HAC}$ that the size and power of the test deteriorates in the presence of inter-unit correlation, as well as in the presence of interunit correlation $\tilde{\Delta}_{CSA}$ test has better performance against autocorrelation, heteroskedasticity as well as inter-unit correlation.

3.2. Eberhardt and Teal (2010) Extended Mean Group (AMG) Estimator

The Alternating Least Squares (ALS) estimator, proposed by Eberhardt and Teal (2010), was developed as an alternative to the Common Component Estimation and Modal Groups (CCEMG) estimator introduced by Pesaran (2006). The ALS estimation process involves three stages.

Initially, the model is considered a one-way time-efficient model, and dummy variables are incorporated into the model for T-1 time periods. The model with time dummies is then estimated using the First Differences Model (FD-OLS), where first differences are taken.

$$\Delta Y_{it} = b'\Delta X_t + \sum_{t=2}^{T} c_t \Delta D_t + e_{it}\bar{F} \Rightarrow> \hat{c}_t = \hat{\mu}_t^* \quad (8)$$

Coefficients of time dummy variables obtained from equation (8) are in a single series $(\hat{\mu}_t^*)$ and these coefficients represent the common dynamic process. In the second stage, the obtained $\hat{\mu}_t^*$ The series is included in the regression equation of each unit with a linear trend:

$$y_{it} = a_i + b_i' x_{it} + c_i t + d_i \hat{\mu}_t^* + e_{it} \quad (9)$$

In the final stage, the coefficients obtained on a per-unit basis are averaged using the Pesaran and Smith (1995) Average Group estimator method, as detailed by Eberhardt (2012: 61-71).

$$\hat{\beta}^{AMG} = \frac{1}{N}\sum_{i=1}^{N} b_i' \quad (10)$$

The findings of Monte Carlo simulations conducted by Eberhardt and Teal (2010) indicate that the AMG estimator exhibits the same asymptotic properties as the CCE estimator.

3.3. Classifier-lasso estimator for heterogenous panel

In the work of Huang et al. (2024), we present a novel approach, the classify lasso method, which employs the c-lasso method to fit panel data models with concealed group structures. This method consists of a post-estimation technique to showcase and visualise the estimated results. There are several ways to enhance the existing classifylasso method. One potential enhancement is to extend the prediction framework to include profile likelihood estimation and the generalised moment estimation approach, as suggested in Su, Shi, and Phillips (2016). Another potential improvement is to increase the speed of the estimation procedure, as computational time can be a concern when working with vast financial datasets. This paper utilises the classifier-lasso method, executed with the classifylasso method, to investigate panel data models by concurrently detecting and estimating unobservable parameter heterogeneity. Experts employ penalised least squares estimation to identify group-specific coefficients and classify unidentified group membership into a predetermined number of groups. This allows for a comprehension of the distinct behaviours exhibited by the various groups in the dataset.

(Initial estimates) The classify lasso method starts with initial (Huang W. et al., 2024)

$$\widehat{\boldsymbol{\alpha}}^{(0)} = \left(\boldsymbol{\alpha}_1^{(0)\prime}, \ldots, \widetilde{\boldsymbol{\alpha}}_K^{(0)\prime}\right)' = \mathbf{0}_{pK\times 1} \text{ and } \widehat{\boldsymbol{\beta}}^{(0)} = \left(\widetilde{\boldsymbol{\beta}}_1^{(0)\prime}, \ldots, \widehat{\boldsymbol{\beta}}_N^{(0)\prime}\right)' \quad (11)$$

Here

$\mathbf{0}_{pK\times 1}$ denotes a $pK \times 1$ zero matrix, and

$$\widetilde{\boldsymbol{\beta}}_i^{(0)} = \begin{cases} (\mathbf{X}_i'\mathbf{X}_i)^{-1}\mathbf{X}_i'\mathbf{y}_i, & \text{if } \mathbf{y}_i'\mathbf{y}_i > 0.0001N \\ (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}, & \text{if } \mathbf{y}_i'\mathbf{y}_i \le 0.0001N \end{cases} \quad (12)$$

where $\mathbf{y}_i$ and $\mathbf{X}_i$ are a $T \times 1$ vector and a $T \times p$ matrix that denote the dependent and independent variables of individual i and $\mathbf{X} = (\mathbf{X}_1', \ldots, \mathbf{X}_N')', \mathbf{y} = (\mathbf{y}_1', \ldots, \mathbf{y}_N')'$. If the variation of the dependent variable of individual i is large enough, the initial value is set to its time-series estimation result; otherwise, the pooled panel one is used.

(Conditional minimization) Suppose we have obtaine d $\widehat{\boldsymbol{\alpha}}^{(r-1)}$ and $\widehat{\boldsymbol{\beta}}^{(r-1)}$ at the r th iteration, $r \geq 1$. Let $Q_{NT}^{\text{ols}}(\boldsymbol{\beta}) = 1/(NT)\sum_{i=1}^{N}\sum_{t=1}^{T}(\tilde{y}_{it} - \boldsymbol{\beta}_i'\tilde{\mathbf{x}}_{it})^2$, and (Huang W. et al., 2024)

$$\begin{cases} \left(\boldsymbol{\beta}^{(r,1)}, \hat{\alpha}_1^{(r)}\right) = \underset{\beta,\alpha_1}{\arg\min}\left(Q_{NT}^{\text{ok}}(\boldsymbol{\beta}) + \frac{\lambda}{N}\sum_{i=1}^{N}\|\boldsymbol{\beta}_i - \boldsymbol{\alpha}_1\|\prod_{k\neq 1}^{K}\left\|\boldsymbol{\beta}_i^{(r-1)} - \hat{\boldsymbol{\alpha}}_k^{(r-1)}\right\|\right) \\ \left(\boldsymbol{\beta}^{(r,2)}, \hat{\alpha}_2^{(r)}\right) = \underset{\beta,\alpha_2}{\arg\min}\left(Q_{NT}^{\text{ok}}(\boldsymbol{\beta}) + \frac{\lambda}{N}\sum_{i=1}^{N}\left\|\boldsymbol{\beta}_i^{(r,1)} - \hat{\boldsymbol{\alpha}}_1^{(r)}\right\|\|\boldsymbol{\beta}_i - \boldsymbol{\alpha}_2\|\prod_{k\neq 1,2}^{K}\left\|\boldsymbol{\beta}_i^{(r-1)} - \hat{\boldsymbol{\alpha}}_k^{(r-1)}\right\|\right) \\ \cdots\cdots \\ \left(\hat{\boldsymbol{\beta}}^{(r,K)}, \hat{\boldsymbol{\alpha}}_K^{(r)}\right) = \underset{\beta,\alpha_K}{\arg\min}\left(Q_{NT}^{\text{olv}}(\boldsymbol{\beta}) + \frac{\lambda}{N}\sum_{i=1}^{N}\prod_{k=1}^{K-1}\left\|\hat{\boldsymbol{\beta}}_i^{(r,k)} - \hat{\boldsymbol{\alpha}}_k^{(r)}\right\|\|\boldsymbol{\beta}_i - \boldsymbol{\alpha}_K\|\right) \end{cases} \quad (13)$$

We hence obtain $\widehat{\boldsymbol{\alpha}}^{(r)} = \left(\widehat{\boldsymbol{\alpha}}_1^{(r)}, \ldots, \widehat{\boldsymbol{\alpha}}_K^{(r)}\right)$ and $\widehat{\boldsymbol{\beta}}^{(r)} = \widehat{\boldsymbol{\beta}}^{(r,K)}$.

(Convergence criterion) The iterative algorithm ends if $r = R_{\max}$ or (Huang W. et al., 2024)

$$\frac{\sum_{i=1}^{N}\left\|\widetilde{\boldsymbol{\beta}}_i^{(r)} - \widetilde{\boldsymbol{\beta}}_i^{(r-1)}\right\|}{\sum_{i=1}^{N}\left\|\widetilde{\boldsymbol{\beta}}_i^{(r)}\right\| + 0.0001} < \epsilon_{\text{tol}} \text{ and } \frac{\sum_{i=1}^{N}\left\|\widetilde{\boldsymbol{\alpha}}_i^{(r)} - \widetilde{\boldsymbol{\alpha}}_i^{(r-1)}\right\|}{\sum_{i=1}^{N}\left\|\widehat{\boldsymbol{\alpha}}_i^{(r)}\right\| + 0.0001} < \epsilon_{\text{tol}} \quad (14)$$

where $R_{\max}$ is the maximum number of iterations and ϵ_{tol} is the tolerance level. Let R be the largest number that meets the above criterion, so $\widehat{\boldsymbol{\beta}} = \widehat{\boldsymbol{\beta}}^{(R)}$ and $\widehat{\boldsymbol{\alpha}} = \widehat{\boldsymbol{\alpha}}^{(R)}$. (Huang W. et al., 2024)

4. Variables, Dataset and Empirical Results

4.1 Variables and dataset

Drawing upon the aforementioned discussion and the pertinent theoretical framework, the present study incorporates three variables that are moderately expected to influence the Consumer Confidence Index (CCI), specifically Artificial Intelligence Tweets on Twitter (AItw), Consumer Price Index (CPI), and Long-term Interest Rates (LTIR).

$$CCI_{it} = f(AItw_{it}, CPI_{it}, LTR_{it}) \quad (15)$$

In Equation (15), CCI_{it} represents the Consumer Confidence Index, $AItw_{it}$ stands for Twitter artificial intelligence tweets, CPI_{it} signifies the Consumer Price Index, and $LTIR_{it}$ refers to the Long-Term Interest Rates. The analysis examined panel data for 28 countries, including 4 non-OECD economies and 24 OECD economies, over a 27-month period from January 2021 to March 2023. The study used the Consumer Confidence Index as the dependent variable and AI tweets on Twitter, Consumer Price Index, and Long-Term Interest Rates as independent variables. The data was obtained from the Organisation for Economic Co-Operation and Development Artificial

Intelligence (OCED AI live data) and Organisation for Economic Co-Operation and Development (OCED.stat) databases. Table 1 provides a comprehensive description of the variables and their definitions.

Table 1: Description of variables and data sources.

Variables	**Symbol**	**Definition**	**Measure**	**Data sources**
OECD Standardised CCI, Amplitude adjusted (Long term average=100), sa	CCI	The OECD Consumer Confidence Index (CCI) is a widely recognized economic indicator that assesses the level of optimism among households with regard to the future economic outlook.	Index	OECD.Stat
X (FORMER TWITTER) artificial intelligence tweets	AItw	Measures the number of hashtags present in tweets containing the terms 'AI' or #ai over time, on a country-by-country basis. The country-specific data is obtained through geotags, which are based on the location specified in the user's profile.	30 day moving average	OECD.AI. Live date
Consumer price indices (CPIs)	CPI	The Consumer Price Index (CPI) of the Organisation for Economic Co-operation and Development (OECD) is a vital gauge of inflation.	%- Percentage change on the same period of the previous year	OECD.Stat
Long-term interest rates	LTIR	The long-term interest rates of the Organisation for Economic Co-operation and Development (OECD) pertain to the interest rates on government bonds that typically have a maturity period of around ten years.	%- Level, ratio or index	OECD.Stat

The present research examines 28 countries, comprising 4 non-OECD and 24 OECD economies, for the period spanning January 2021 to March 2023. The rationale underpinning the choice of

these countries is that there is no loss of data. Table 2 presents an summary of the countries categorised by group.

Table 2: Selected countries

OECD Economies		Non-OECD Economies
Austria	Israel	Brazil
Belgium	Italy	China
Chile	Korea	Indonesia
Colombia	Luxembourg	South Africa
Czech Republic	Netherlands	
Denmark	Poland	
Estonia	Portugal	
Finland	Spain	
France	Sweden	
Germany	Switzerland	
Greece	United Kingdom	
Hungary	United States	

4.2 Econometric model

Prior to undertaking the analysis, in order to simplify the intricate nature of equation (15), we transformed the variables CCI_{it} - Consumer confidence index and $AItw_{it}$ - Artificial intelligence tweets on Twitter into logarithmic forms, thereby enhancing the lucidity of the outcomes.

$$\ln CCI_{it} = \beta_0 + \beta_1 \ln AItw_{it} + \beta_2 CPI_{it} + \beta_3 LTIR_{it} + \varepsilon_{it} \quad (16)$$

The heterogeneous panel data model is widely employed to assess the impact of X (Old Twitter) artificial intelligence tweets on the consumer confidence index across a group of selected countries, with monthly data spanning the period January 2021 to March 2023.

In order to assess the causality of the endogenous variables in our model, we employ the panel causality test for panel Granger causality developed by Xiao et al. (2023) and Juodis et al. (2021), which accounts for horizontal cross-section dependence for both homogeneous and heterogeneous coefficients. The p-values for the Wald statistic indicate the presence of unidirectional and bidirectional Granger causality, as demonstrated by Guliyev.H. (2022) in Table 3.

Table 3. Panel causality test results.

Panel A: Multivariate non-causality test					
H_0: Selected covariates do not Granger-cause lnCCİ	**Lags**	**HPJ Wald test**	**p-value**	**Decision**	**Conclusion**

lnAItw, CPI, LTIR⇒ EFC	1	54.585*	0.00	H_0 is reject	AI tweets on Twitter are Granger-caused lnCCI
Panel B: Univariate non-causality test					
H_0: x does not Granger-cause (⇒) y	**Lags**	**HPJ Wald test**	**p-value**	**Decision**	**Conclusion**
lnAItw ⇒ lnCCI	1	4.615*	0.0317	H_0 is reject	Bidirectional
lnCCI⇒ lnAItw	1	5.783*	0.0162	H_0 is reject	lnCCI ⇔ lnAItw
CPI ⇒ lnCCI	4	12.885*	0.0118	H_0 is reject	Bidirectional
lnCCI ⇒ CPI	1	17.180*	0.00	H_0 is reject	lnCCI ⇔ CPI
LTIR⇒ lnCCI	3	116.281*	0.00	H_0 is reject	Unidirectional
lnCCI ⇒ LTIR	1	1.571*	0.21	H_0 is not rejected	LTIR⇒ lnCCI

Note: ∗∗∗, ∗∗, and ∗ denote statistical significance at the 1%, 5%, and 10% levels, respectively.

Before selecting an estimation method for heterogeneous panel data models, it is essential to examine various features of the panel data structure. We do this to ensure we use the correct specification when determining the tests and estimators to employ in the subsequent stage. In Table 4, the first step involves assessing whether the individual variables and, subsequently, the model errors exhibit inter-unit correlation.

Table 4: Assessment of the Slope Parameter's Homogeneity in the Panel Data Model

	Swamy (1970) S testi	Pesaran ve Yamagata (2008) **Δ̃**Testi	Pesaran ve Yamagata (2008) **Δ̃** adj. Testi	Blomquist ve Westerlund (2013) **Δ̃** HAC testi	Blomquist ve Westerlund (2013) **Δ̃** adj. HAC testi	Bersvendsen and Ditzen (2021) **Δ̃** CSA testi	Bersvendsen and Ditzen (2021) **Δ̃**CSA adj testi
Test statistic	2996.02	19.09	21.53	2.40	2.70	19.09	21.53
Prob	0.00	0.00	0.00	0.016	0.007	0.00	0.00

For $\text{Delta}_{\text{Hac}}$, $\text{Delta}_{\text{Hac-adj}}$ $\text{Delta}_{\boldsymbol{CSA}}$ and Delta $_{\boldsymbol{CSA\text{-}addj}}$ tests, the QS (Quadratic-Spectral) function was used as the Kernel function to obtain the HAC standard errors to control heteroskedasticity and autocorrelation.

In Delta $_{\boldsymbol{CSA}}$ and Delta $_{\boldsymbol{CSSA\text{-}addj}}$ tests, horizontal cross-section averages of the variables were used in the test process to control the correlation between units.

Note: ∗∗∗, ∗∗, and ∗ denote statistical significance at the 1%, 5%, and 10% levels, respectively.

In Table 4, Swamy (1970) employed the $\tilde{S}$ test, while Pesaran and Yamagata (2008) used the $\tilde{\Delta}$ test and its small sample bias-corrected version, $\tilde{\Delta}adj$. These tests were used to examine the homogeneity of slope coefficients in the panel data model. Blomquist and Westerlund (2013) utilized the $\tilde{\Delta}HAC$ test and its small sample corrected version, $\tilde{\Delta}HAC\text{-}adj$, which are robust to heteroskedasticity and autocorrelation. Lastly, Bersvendsen and Ditzen (2021) employed the $\tilde{\Delta}$ CSA test and its small sample corrected version, $\tilde{\Delta}CCSA\text{-}adj$, which are robust to heteroskedasticity, autocorrelation, and inter-unit correlation.

According to the findings of Swamy (1970) $\tilde{S}$, Pesaran and Yamagata (2008) $\tilde{\Delta}$ and $\tilde{\Delta}adj$ tests, the null hypothesis for the homogeneity of slope parameters is dismissed at the 5% confidence level, indicating that slope coefficients differ from unit to unit, i.e. they are heterogeneous. Nevertheless, these tests are not resistant to potential heteroskedasticity and autocorrelation, as indicated by the literature. Therefore, the $\tilde{\Delta}$ HAC and $\tilde{\Delta}HAC\text{-}adj$ tests suggested by Blomquist and Westerlund (2013) are more dependable. As a result of these tests, the null hypothesis that the slope coefficients are homogeneous is rejected at the 5% significance level, and it is concluded that the slope coefficients vary from unit to unit, i.e. they are heterogeneous. Bersvendsen and Ditzen (2021) argue in their study that if the variables exhibit strong inter-unit correlation and autocorrelation, extending the $\tilde{\Delta}HAC$ test with the horizontal cross-sectional averages of the variables and obtaining the $\tilde{\Delta}CSA$ test enhances the power and size properties of the test. Considering Tables 4 and 5, it is demonstrated that the error term of the model has inter-unit correlation in semi-strong form, thus it is more appropriate to utilize the $\tilde{\Delta}CSA$ and $\tilde{\Delta}CSA\text{-}adj$ test, which takes into account all assumption violations. According to the findings of this test, the null hypothesis testing the homogeneity of the slope parameters is dismissed at the 5% confidence level, indicating that the slope coefficients vary from unit to unit, and the panel data model is deemed to be heterogeneous.

Table 5: Extended Mean Group (AMG) estimator

Variables	**Coef.**	**Std. errors**	**prob.**
lnAItw	0.009766	0.004283	0.023
CPI	-0.00297	0.00098	0.002
LTIR	0.010124	0.001907	0.000
Constant	4.500011	0.041755	0.000
N=28 T=27 **CDw+=969.99 prob=0.00** CD*=**1.92 prob=0.055** **Wald chi2(3) = 75.8 (prob<0.01)** Eberhardt and Teal (2010), Eberhardt and Bond (2009) derived Extended Mean Group (AMG) estimator for the joint process by taking μˆ_t*ʾ as the difference from the dependent variable			

The outcomes of the Pesaran (2004, 2015) CD test, as well as the CDw test developed by Juodis and Reese (2022), the CDw+ test extended with the power enhancement approach of Fan et al. (2015), and the CD* test extended with the factors of Pesaran and Xie (2022) are presented in Table 5. When selecting an inter-unit correlation test, the size of the panel should be taken into account. Considering that N=28 and T=27 in the data set used, and that T is smaller than N, the results of the LM adj*- Pesaran et al. (2008) Bias-adjusted scaled Lm adj* and Pesaran (2004, 2015) CD test, and the CDw test developed by Juodis and Reese (2022) and the CDw+ test extended with the power enhancement approach of Fan et al. (2015) and the CD* test extended with the factors of Pesaran and Xie (2022) are more reliable. All the tests performed, as well as the test results of LM adj*- Pesaran et al. (2008) Bias-adjusted scaled Lm adj* and Pesaran (2004,2015) CD test, and CDw test developed by Juodis and Reese (2022) and CDw+ test extended with the power enhancement approach of Fan et al. (2015), reject the null hypothesis of no correlation between units at the 5% significance level for all variables.

Moreover, when the α values obtained from the Pesaran (2015) CD test are taken into account, it is concluded that there is a strong inter-unit correlation for these variables since the α values for the Consumer confidence index, X (Old Twitter) artificial intelligence tweets, Consumer price index, and Long-term interest rates variables are approximately one.

Table 5 presents the results of the estimation and testing of assumptions for the EMGAMG joint process, which involves differencing $\hat{\mu}_t^*$ from the dependent variable. The table showcases the extended co-correlated effects estimator along with the Wald test statistic and the corresponding probabilities. The results of the Wald test indicate that the model is statistically significant. All variables in the Extended Mean Group (AMG) estimator, derived by Eberhardt and Teal (2010) and Eberhardt and Bond (2009) by taking the difference of $\hat{\mu}_t^*$ from the dependent variable for the joint process, are significant at the 1%, 5%, and 10% significance levels. The Z statistic for the significance of the regression models for each individual country varies across countries and will be discussed in detail in the results section of the paper.

Table 6: Classifier-lasso Estimator Results

Group 1	**Group 2**
Austria	Chile
Belgium	China (People's Republic of)
Brazil	Denmark
Colombia	Estonia
Czechia	Finland
Germany	France
Greece	Israel
Hungary	Korea
Indonesia	Luxembourg
Italy	Poland
Netherlands	Portugal

Spain United States	South Africa Sweden Switzerland United Kingdom

Figure 2: Illustration of the implementation instance

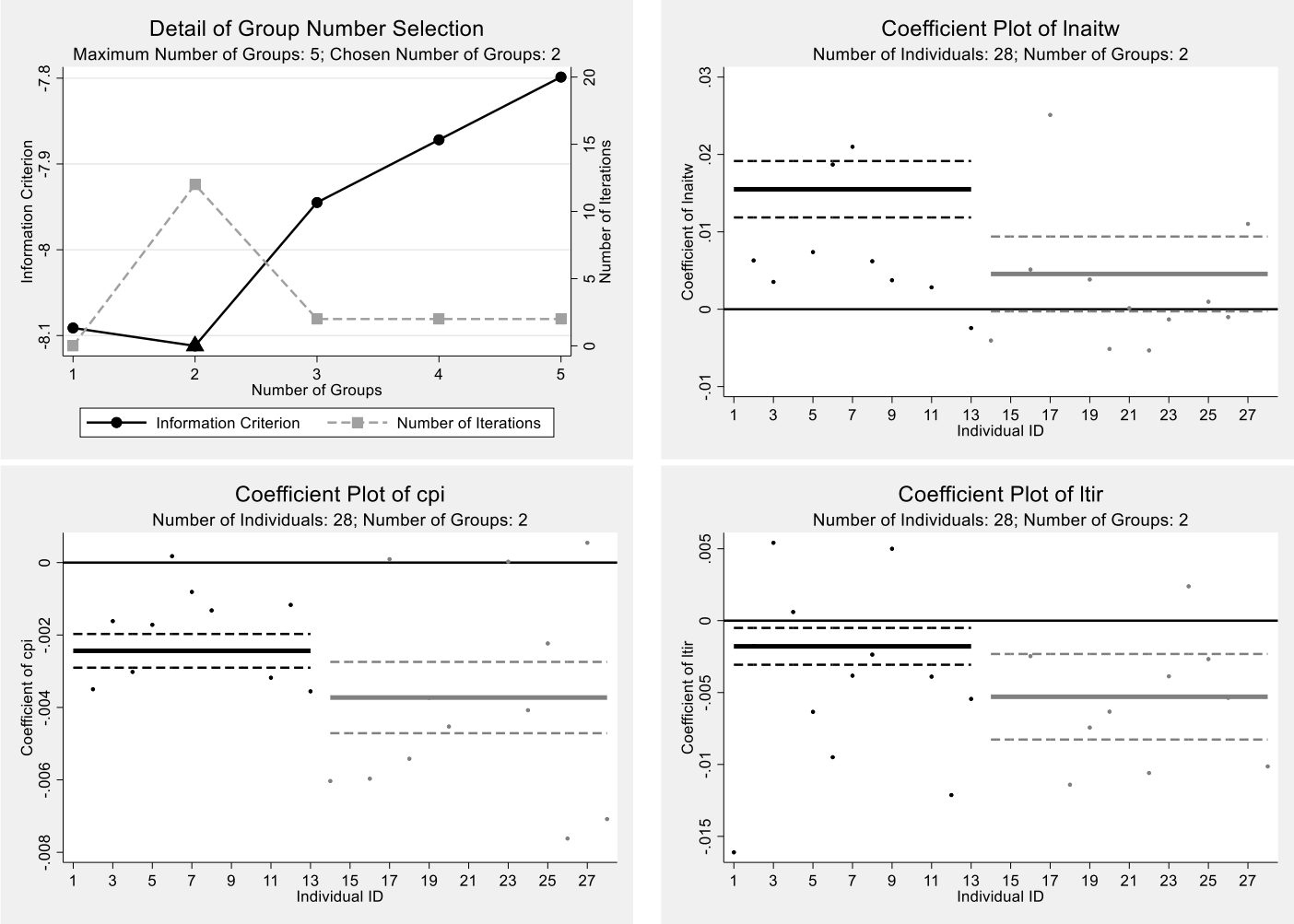


Figure 2 (left) depicts the number of groups selection information. The black dots represent the values of the information criterion in relation to the number of groups. The triangle in the graph indicates the number of groups selected, and the minimum information criterion is marked by this triangle. The grey dots indicate the number of iterations used. In Figure 2 (right), the coefficient estimates of lnaitw, cpi, and ltri are illustrated. Different shades are used to represent different groups. Solid and short lines plot the group-specific postlasso estimates and confidence bands, while dots display the individual-specific time series point estimates with outliers removed.

Table 7: Group 1 Estimation Results

Variables	FE		RE	
	coef	**Driscall-Kraay standart error**	**coef**	**Driscall-Kraay standart error**
lnAItw	.0155018***	.0018592	.0056874***	.0013762
CPI	-.0024364***	.0002375	-.0024313***	.0002314
LTIR	-.0017845***	.0006534	-.0013832**	.0006026
Constant	4.479334***	.0163902	4.564457***	.0126126
N= 28 T=27 Hausman test - chi2(3) = 63.58, Prob > chi2 = 0.0000 Heteroskedasticity test (Modified Wald test for groupwise heteroskedasticity)- chi2 (13) = 243.50, Prob>chi2 = 0.0000 Autocorrelation test - Modified Bhargava et al. Durbin–Watson = 2.7826394, Baltagi–Wu LBI = 2.8215496 Inter-unit correlation test - Pesaran's test of cross sectional independence = 12.163, Pr = 0.0000, Friedman's test of cross sectional independence = 89.998, Pr = 0.0000 Note: ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.				

Table 8. Group 2 Estimation Results

Variables	FE		RE	
	coef	**Driscall-Kraay standart error**	**coef**	**Driscall-Kraay standart error**
lnAItw	.0045575*	.0024665	-.0015026	.0013642
CPI	-.0037266***	.0005023	-.0043601***	.0003473
LTIR	-.0052928***	.0015181	-.0012655	.0008073
Constant	4.585237***	.0198173	4.627878***	.0113002
N= 28 T=27 Hausman test - chi2(3) = 53.17, Prob > chi2 = 0.0000 Heteroskedasticity test (Modified Wald test for groupwise heteroskedasticity)- chi2 (15) = 3211.38, Prob>chi2 = 0.0000 Autocorrelation test - Modified Bhargava et al. Durbin–Watson = 3.3021353, Baltagi–Wu LBI = 3.3521723 Pesaran's test of cross sectional independence CD= 17.347, Pr = 0.0000, Friedman's test of cross sectional independence = 131.547, Pr = 0.0000 Note: ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.				

Table 7-8 exhibits the estimation results and assumption tests for the Classifier-lasso method (Su, Shi, and Phillips, 2016). The primary objective of this technique is to detect unobserved parameter heterogeneity in panel data models using penalised techniques and to determine group-specific coefficients, classify unknown group membership, and establish the number of groups based on an information criterion.

The table indicates that the regression equations for Group 1's specific countries, which encompass Old Twitter AI tweets (AItw), Consumer Price Index (CPI), and Long Term Interest Rates (LTIR), provide a medium-high level of explanation for the Consumer Confidence Index (CCI). In the classifier-lasso estimation, all variables in Group 1 were found to be statistically significant at the 1%, 5%, and 10% levels of significance. The Z statistic computed for the significance of the regression models of the countries in the group was also found to be statistically significant.

5. Conclusion and policy implications

This study investigates the disparities between the Consumer Confidence Index (CCI) and Twitter discussions on artificial intelligence (new name X), the Consumer Price Index (CPI), and Long-term Interest Rates (LTIR) across 24 Organization for Economic Co-operation and Development (OCED) countries and 4 non-OCED countries for a period of 27 months, from January 2021 to March 2023. The primary variable under examination is the Consumer Confidence Index (CCI), which gauges consumers' perceptions of their financial circumstances and their expectations for the nation's economic climate, as well as their tendencies to save and spend in the near future. Fluctuations in the CCI can have a profound impact on macroeconomic indicators and the overall economic health of a country.

To estimate the heterogeneous panel econometric model, we follow the process outlined in Figure 1 (Flowchart summarising the analysis of the heterogeneous panel model).

In continuation of the flowchart, Swamy's (1970) $\tilde{S}$ Test, Pesaran and Yamagata's (2008) $\Delta\tilde{}$ Test, Blomquist and Westerlund's (2013) $\tilde{\Delta}$ test is employed to determine the appropriate econometric model between Homogeneous and Heterogeneous Panel Data Models. Additionally, Bersvendsen and Ditzen's (2021) χ_CSA test is utilized in the analysis. The outcomes of the tests indicate that the null hypothesis, which assumes the homogeneity of the slope parameters, is rejected at the 5% confidence level. This suggests that the slope coefficients vary across different units, and therefore, the panel data model is heterogeneous in nature. No changes to the citation or reference style are permitted.

Breusch-Pagan (1980) LM Test, Pesaran (2004, 2015) CD Test, Pesaran, Ullah and Yamagata (2008) Adjusted LM Test, Juodis and Reese (2022) CD Test, Pesaran and Xie (2022) CD^* Upon evaluating the test results, the null hypothesis of no correlation between units is rejected at a 5% significance level for all variables. Furthermore, when the alpha values obtained from the Pesaran (2015) CD test are taken into consideration, it is determined that there is a strong inter-unit correlation for these variables, as the alpha values for the Consumer Confidence Index, X (Old Twitter) artificial intelligence tweets, Consumer Price Index, and Long-term interest rates variables are approximately equal to one. In general, it is concluded that the analysed panel structure is both

correlated between units and heterogeneous according to the slope coefficients. In this case, in the presence of inter-unit correlation, it is appropriate to use heterogeneous estimators that take into account the inter-unit correlation. These estimators control for inter-unit correlation with the help of common factors.

In the subsequent phase, the Extended Mean Group (AMG) method, which was derived by Eberhardt and Teal (2010) and Eberhardt and Bond (2009) for the estimation of panel data models by subtracting μ̂_t* from the dependent variable and employing penalized techniques, was used to simultaneously identify unobserved parameter heterogeneity in panel data models and classifier-lasso tests for estimation (Su, Shi, and Phillips, 2016, Econometrica 84: 2215-2264).

According to the results of the AMG estimation, all variables were found to be statistically significant. Specifically, a one-unit increase in Artificial Intelligence Tweets was associated with a 0.0097 unit increase in the Consumer Confidence Index. Furthermore, a 1% increase in the Consumer Price Index was associated with a 0.297% decrease in the Consumer Confidence Index. Additionally, a 1% increase in long-term interest rates was associated with a 1.0124% increase in the Consumer Confidence Index.

According to Su, Shi, and Phillips (2016), the classifier-lasso method revealed the importance of all variables in the estimation results for the 13 selected countries. A rise of one unit in Artificial Intelligence Tweets resulted in a 0.015 unit increase in the Consumer Confidence Index, while a 1% increase in the Consumer Price Index led to a decrease of -0.244% in the Consumer Confidence Index. Furthermore, a 1% rise in long-term interest rates resulted in a decrease of -0.178% in the Consumer Confidence Index.

In the second group of fifteen countries, where all variables were statistically significant, an increase of one unit in Artificial Intelligence Tweets was correlated with a 0.0045 unit rise in the Consumer Confidence Index. Additionally, a 1% increase in the Consumer Price Index was associated with a decrease of -0.373% in the Consumer Confidence Index. Lastly, a 1% rise in long-term interest rates led to a decrease of -0.529% in the Consumer Confidence Index.

Social media's growing significance in the economy has led to the recognition of the need for consumers, as economic participants, to monitor and analyse social media data to make well-informed decisions and predict economic trends. In general, Twitter discussions on AI represent the changing perceptions of the role of technology in shaping consumer behaviour and expectations, with a balance of positive sentiments and concerns about the impact of AI on the economy and society. It is possible that tweets about AI may influence consumers' perceptions of the economy and their own financial circumstances. Specifically, the potential impact of AI technologies on labour markets may change consumers' economic expectations for the future. Positive or negative perceptions of AI may also reflect consumers' spending and saving tendencies. Positive perceptions may lead consumers to feel more secure and increase their spending, while negative perceptions may cause consumers to be more cautious and reduce their spending. In conclusion, the content and tone of tweets about artificial intelligence can be considered an important factor that may affect the consumer confidence index. However, other macroeconomic and social factors that affect consumer confidence should also be taken into account.

Funding

This study was not funded.

Declaration of interests

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

Acemoglu, Daron, and Andrew Scott. "Consumer confidence and rational expectations: Are agents' beliefs consistent with the theory?." The Economic Journal 104.422 (1994): 1-19.

Ante, Lennart. 2023. “How Elon Musk’s Twitter Activity Moves Cryptocurrency Markets.” Technological Forecasting and Social Change 186 (January): 122112.

Bai, J., & Ng, S. (2004). A PANIC attack on unit roots and cointegration. Econometrica, 72(4), 1127-1177.

Bai, J., & Ng, S. (2010). Panel unit root tests with cross-section dependence: a further investigation. Econometric Theory, 26(4), 1088-1114.

Baltagi, B. H. (Ed.). (2015). The Oxford handbook of panel data. Oxford University Press.

Baltagi, Badi H. (2021) Econometric analysis of panel data, Springer, 6. bs,

Bersvendsen, T., & Ditzen, J. (2021). Testing for slope heterogeneity in Stata. The Stata Journal, 21(1), 51-80.

Bianchi, F., Gómez-Cram, R., Kind, T., & Kung, H. (2023). Threats to central bank independence: High-frequency identification with twitter. Journal of Monetary Economics, 135, 37-54.

Bianchi, Francesco, Roberto Gómez-Cram, and Howard Kung. "Using social media to identify the effects of congressional viewpoints on asset prices." The Review of Financial Studies 37.7 (2024): 2244-2272.

Bodaghi, Amirhosein, and Jonathan JH Zhu. "A big data analysis of the adoption of quoting encouragement policy on Twitter during the 2020 US presidential election." Journal of Computational Social Science (2024): 1-33.

Breusch, T.S., and A.R. Pagan (1980): “The Lagrange Multiplier Test and its Applications to Model Specification in Econometrics,” Review of Economic Studies 47, 239-253.

Cao, Andy, Jason M. Lindo, and Jiee Zhong. "Can social media rhetoric incite hate incidents? Evidence from Trump's “Chinese Virus” tweets." Journal of Urban Economics 137 (2023): 103590.

Carroll, C. D., Fuhrer, J. C., & Wilcox, D. W. (1994). Does consumer sentiment forecast household spending? If so, why?. The American Economic Review, 84(5), 1397-1408.

Christ, Kevin P., and Dale S. Bremmer. "The relationship between consumer sentiment and stock prices." New York Times (2001).

Chudik, A., & Pesaran, M. H. (2015). Common correlated effects estimation of heterogeneous dynamic panel data models with weakly exogenous regressors. Journal of econometrics, 188(2), 393-420.

Chudik, A., Pesaran, M. H., & Tosetti, E. (2011). Weak and strong cross-section dependence and estimation of large panels.

Cookson, J. Anthony, William Mullins, and Marina Niessner. "Social Media and Finance." Available at SSRN 4806692 (2024).

Corina, Pelau., Irina, Ene., Mihai-Ionut, Pop. (2021). The Impact of Artificial Intelligence on Consumers Identity and Human Skills. Amfiteatru Economic, doi: 10.24818/EA/2021/56/33

De Hoyos, R. E., & Sarafidis, V. (2006). Testing for cross-sectional dependence in panel-data models. The stata journal, 6(4), 482-496.

De Witte, H., Pienaar, J., & De Cuyper, N. (2016). Review of 30 years of longitudinal studies on the association between job insecurity and health and well-being: Is there causal evidence?. Australian Psychologist, 51(1), 18-31.

DEMİRÖZ, D. M. (2019). Technological and Socio-Economic Effects of Fifth and Sixth Kondratieff Waves. Istanbul Journal of Economics= İstanbul İktisat Dergisi, 69(2), 247.

Eberhardt, M., & Bond, S. (2009). Cross-section dependence in nonstationary panel models: a novel estimator.

Eberhardt, Markus and Teal, Francis (2010). "Econometrics for Grumblers: A New Look at the Literature on Cross-Country Growth Empirics." Journal of Economic Surveys. Forthcoming

Erlinghagen, M. (2008). Self-perceived job insecurity and social context: A multi-level analysis of 17 European countries. European Sociological Review, 24(2), 183-197.

Guliyev, H. (2022). The effect of global financial markets and local shocks on Turkey airlines market; new evidence from structural break cointegration and causality tests. Research in Globalization, 5, 100096.

Garrett, T. A., Hernández-Murillo, R., & Owyang, M. T. (2005). Does consumer sentiment predict regional consumption. Federal reserve Bank of st. Louis review, 87(2), 123-135.

Golinelli, R., & Parigi, G. (2004). Consumer sentiment and economic activity: a cross country comparison. Journal of Business Cycle Measurement and Analysis, 2004(2), 147-170.

Gujarati,Damodar N., Porter, Dawn C. (2012) Basic Econometrics, McGraw-Hill Education, 5.bs,

Huang, W., Wang, Y., & Zhou, L. (2024). Identify latent group structures in panel data: The classifylasso command. The Stata Journal, 24(1), 46-71.

J. Fan, Y. Liao, J. Yao, Power enhancement in high-dimensional cross-sectional tests, Econometrica 83 (4) (2015) 1497–1541.

Jain, Varsha, Ketan Wadhwani, and Jacqueline K. Eastman. "Artificial intelligence consumer behavior: A hybrid review and research agenda." Journal of consumer behaviour 23.2 (2024): 676-697.

Juodis, A., & Reese, S. (2022). The incidental parameters problem in testing for remaining cross-section correlation. Journal of Business & Economic Statistics, 40(3), 1191-1203.

Katona, George, "Consumer Behavior: Theory And Findings On Expectations And Aspirations" The American Economic Review, Vol. 58, No. 2, May 1968, ss. 19-30.

Kuhnen, C. M., & Knutson, B. (2011). The influence of affect on beliefs, preferences, and financial decisions. Journal of Financial and Quantitative Analysis, 46(3), 605-626.

Linwan, Wu., Naa, Amponsah, Dodoo., Taylor, Jing, Wen., Li, Ke. (2021). Understanding Twitter conversations about artificial intelligence in advertising based on natural language processing. International Journal of Advertising, doi: 10.1080/02650487.2021.1920218

Masciandaro, Donato, Oana Peia, and Davide Romelli. "Central bank communication and social media: From silence to Twitter." Journal of Economic Surveys 38.2 (2024): 365-388.

Mehra, Y. P., & Martin, E. (2003). Why does consumer sentiment predict household spending?. FRB Richmond Economic Quarterly, 89(4), 51-67.

Millet, Barbara, et al. "In twitter we trust (ed): How perceptions of Twitter's helpfulness influence news post credibility perceptions and news engagement." Computers in Human Behavior (2024): 108185.

Miyazaki, K., Murayama, T., Uchiba, T., An, J., & Kwak, H. (2024). Public perception of generative AI on Twitter: an empirical study based on occupation and usage. EPJ Data Science, 13(1), 2.

Molla, Rani. 2021. "Chart: How Bitcoin Prices Move with Elon Musk's Tweets." Vox. May 18, 2021. https://www.vox.com/recode/2021/5/18/22441831/elon-musk-bitcoin-dogecoin-crypto-prices-tesla.

Olan, Femi, et al. "Advancing consumer behavior: The role of artificial intelligence technologies and knowledge sharing." IEEE Transactions on Engineering Management (2021).

Pesaran, M. H. (2007). A simple panel unit root test in the presence of cross-section dependence. Journal of applied econometrics, 22(2), 265-312.

Pesaran, M. H. (2015). Testing weak cross-sectional dependence in large panels. Econometric reviews, 34(6-10), 1089-1117.

Pesaran, M. H. (2015). Testing weak cross-sectional dependence in large panels. Econometric reviews, 34(6-10), 1089-1117.

Pesaran, M. H., & Yamagata, T. (2008). Testing slope homogeneity in large panels. Journal of econometrics, 142(1), 50-93.

Pesaran, M. H., Ullah, A., & Yamagata, T. (2008). A bias-adjusted LM test of error cross-section independence. The econometrics journal, 11(1), 105-127.

Pesaran, MH, Smith, LV ve Yamagata, T. (2013). Çok faktörlü hata yapısının varlığında panel birim kök testleri. Ekonometri Dergisi , 175 (2), 94-115.

Phillips, P. C., & Moon, H. R. (1999). Linear regression limit theory for nonstationary panel data. Econometrica, 67(5), 1057-1111.

Reese, S., & Westerlund, J. (2016). Panicca: Panic on cross-section averages. Journal of Applied Econometrics, 31(6), 961-981.

Shayaa, S., Ainin, S., Jaafar, N. I., Zakaria, S. B., Phoong, S. W., Yeong, W. C., ... & Zahid Piprani, A. (2018). Linking consumer confidence index and social media sentiment analysis. Cogent Business & Management, 5(1), 1509424.

Su, L., Shi, Z., & Phillips, P. C. (2016). Identifying latent structures in panel data. Econometrica, 84(6), 2215-2264.

Swamy, P. A. (1970). Efficient inference in a random coefficient regression model. Econometrica: Journal of the Econometric Society, 311-323.

Tahsin, E., & Börü, F. (2023). TÜRKİYE'NİN ENDÜSTRİ 4.0'A GEÇİŞ SÜRECİNİN TEKNOLOJİ KAPASİTESİ YAKLAŞIMLARI TEMELLİ ANALİZİ. Abant Sosyal Bilimler Dergisi, 23(3), 1293-1314.

Vuchelen, J. (2004). Consumer sentiment and macroeconomic forecasts. Journal of economic psychology, 25(4), 493-506.

Westerlund, J., & Blomquist, J. (2013). A modified LLC panel unit root test of the PPP hypothesis. Empirical economics, 44, 833-860.

Westerlund, J., & Blomquist, J. (2013). A modified LLC panel unit root test of the PPP hypothesis. Empirical economics, 44, 833-860.

Xie, Y., & Pesaran, M. H. (2022). A Bias-Corrected Cd Test for Error Cross-Sectional Dependence in Panel Data Models with Latent Factors. Available at SSRN 4198155.

Yerdelen Tatoğlu, F. (2020). Panel Veri Ekonometrisi Stata Uygulamalı. İstanbul: Beta Basım Yayım Dağıtım A.Ş.

Yerdelen Tatoğlu, F. (2021). Advanced Panel Data Analysis Stata Applied. 4th Edition. Istanbul: Beta Basım Yayım Dağıtım