A Dynamic Model of Customers Behavior: Integrating Econophysics and Physics-Informed Neural Networks

Kirill Zakharov[1,*][0000-0001-5774-4076], Anton Kovantsev[1][0000-0003-1765-7001], and Alexander Boukhanovsky[1][0000-0003-1588-8164]

1 ITMO University, Kronverksky Pr. 49, Saint Petersburg, 197101, Russia
Correspondence: kazakharov@itmo.ru

Abstract. The transactional activity of a bank's customers contains a wealth of information regarding their behavior. By examining the transactions of either a specific group or all clients of the bank, it is possible to gain insight into the macroeconomic environment and utilize this to anticipate various outcomes. In this paper, we proposed a dynamic behavioral model for bank clients based on econophysics principles. Additionally, we identified the non-homogeneous function in the dynamic equation using physics-informed neural network. We also interpreted this non-homogeneity through the lens of news reports from social media platforms and news agencies. The model demonstrated accurate results in a numerical simulation of the restoration of the initial dependency. We also showed the potential for creating scenarios in which news events impact the behavior of bank customers.

Keywords: econophysics · Lagrangian mechanics · physics-informed neural networks · customers behaviour · bank transactions.

1 Introduction

In banks or other financial organisations that handle customer transactions, it is essential to understand customer behavior in order to anticipate certain scenarios and enhance the user experience. Transactions contain a substantial amount of data, enabling the provision of insights into the historical patterns of customer behavior and potential purchasing preferences within specific product or service categories.

The primary applications of transactional data encompass: forecasting expenditures [1] and cashback categories [2], providing personalized recommendations [3], detecting fraudulent activities [4], implementing privacy-preserving mechanisms [5], analyzing advertising impact [6], and managing risks [7], among others. All these applications are widely used in modern banking and finance organizations to enhance organizational profitability and analyze potential investment opportunities.

In our article, the objective was to describe the dynamics of customer spending. Specifically, we were interested in a particular feature that denotes the

ICCS Camera Ready Version 2025
To cite this paper please use the final published version:
DOI: 10.1007/978-3-031-97626-1_17

amount of money utilized in the current transaction. This amount of spending is a continuous variable and, furthermore, strictly positive. By understanding the spending patterns of a bank's customers, it is possible to gain insight into the state of the economy, including the presence of potential crisis situations, holidays or other significant events. For instance, the paper [8] explores the recognition of the behavior and experiences of individuals during crisis situations. The work [9] specifically focuses on bank customers and their transactional activity.

Additionally, it is important to understand how bank customers react to external economic factors such as politics, macroeconomic indicators and news. This understanding provides government influence to essential factors that can increase or decrease the purchasing power through the imposition of restrictions, prohibitions, or other regulatory measures.

The description of economic entities and agents is approached through various methods, including the use of various mathematical theories. For example, agent-based modelling, cellular automata, gravitational models. Therefore, studies that utilize a priori economic knowledge and specific physical techniques to construct models of consumption and production [10], social group interactions [11], and other social aspects are of particular interest. However, due to the vast number of interactions between economic agents and other external factors, it is not possible to fully describe behavioural patterns.

We used the tools of dynamical systems theory to identify the behavioural patterns. Specifically, we are concerned with second-order differential equations, as they are relevant to an approach originating from econophysics. This method combines advanced physical principles with economic phenomena through the interpretation of economic indicators using physical analogies.

In econophysics the economic entities, such as production volumes, are expressed in terms of generalized coordinates, and kinetic and potential energies are determined for them. By incorporating the necessary components into the potential energy, production processes can be understood in various ways. For example, potential energy can be defined in such a way that it represents an opportunity for the company to produce more output. This accumulated effect can be realized by increasing the company's utility function [10].

Transactions typically comprise a series of both continuous and discrete factors. These factors may include, for instance, the age of the customers, the amount spent, the transaction category, the terminal address, the customer identification number, the date and time of the transaction.

The process of receiving transactions is a dynamic process, although the receipt of transactions occurs at discrete points in time and at irregular intervals. However, it is more convenient to view transactions as a continuous stream and consider discrete points of time as defined by historical data. For instance, categorical features can be encoded to numerical values, as demonstrated in [12].

Numerous research papers are focused on predicting customer behavior using machine learning techniques. Understanding customer behavior can lead to enhanced company processes and improved customer experiences, as evidenced by studies such as [13] and [14]. Additionally, works related to customer purchases

ICCS Camera Ready Version 2025
To cite this paper please use the final published version:
DOI: 10.1007/978-3-031-97626-1_17

and digital marketing strategies, such as [15] and [16], demonstrate the potential of predicting customer behavior based on transaction data using numerical and categorical attributes. However, these models primarily focus on identifying patterns in the data without fully explaining the underlying causes.

Therefore, despite some work done in the area of modelling the behaviour of bank customers, models are either based on a predetermined behavioural model or on a specific set of features, with further identification of non-linear relationships between these features and the target variable. The solution may involve identifying the underlying forces, the external field, which help to explain the discrepancies between empirical observations and theoretical predictions.

Our contribution:

i) We developed a dynamic model that describes customer expenditures using Lagrangian mechanics. This derivation is based on the kinetic and potential energy associated with the customers behavior.
ii) We identified the non-homogeneous function in the dynamic equation utilizing a physics-informed neural network.
iii) We interpreted the non-homogeneous function by utilizing information from social media and news agencies. We also created scenarios based on news in order to alter the non-homogeneity function and, therefore, change the dynamics of the system.

2 Methods

2.1 Problem formulation

We consider customer behaviour through their bank transactions. Each client has their own views and unique experiences that determine their actions. However, their behavior may exhibit some periodic patterns and seasonal variations, as well as some characteristics that deviate from conventional patterns due to some external factors.

We classified the transaction data into three categories by MCCs (Merchant category codes). The first category consists of all codes related to personal development, *self-realization*. They may include educational pursuits, namely university studies, vocational training courses, musical academies, or specialized master classes. The second category pertains to *socialization* activities. These encompass outings to restaurants, museums, or exhibitions. It can also involve spendings for cinemas, shopping malls, and travel-related activities [17].

The final category of goods and services encompasses everything that individuals require for *survival*. Specifically, pharmacies, hospitals, food and beverages, clothing, utility payments, etc.

Figure 1 shows the client expenses. We have presented the initial data and also constructed a smoothing trend using a period of 14 days and a one-sided filter. The Figure shows, some distinct peaks for the New Year event, corresponding to an increase in the purchasing power. Furthermore, there has been

ICCS Camera Ready Version 2025
To cite this paper please use the final published version:
DOI: 10.1007/978-3-031-97626-1_17

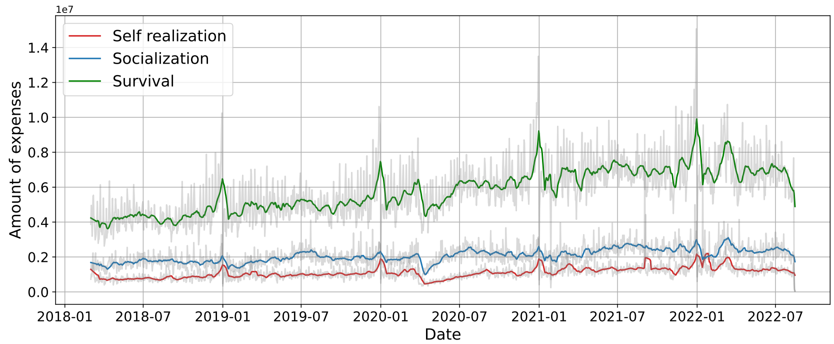


Fig. 1: Customer expenses. The X-axis shows the date, and the Y-axis indicates the total amount spent. The initial data are indicated in gray, while their smoothed trends are represented in color.

an increase in the all three categories, indicating that people are actively involved in events, buying gifts, and purchasing food for the holidays.

It is remarkable that the fairly stable amount of spending in the categories of self-realization and socialization, in contrast to survival, where stable growth occurs, with the exception of crisis periods. We also observed a substantial reduction in expenditures during the first six months of 2020, which coincides with the onset of the COVID-19 pandemic. Furthermore, there was a significant shock in the early periods of 2022 due to a special military operation in Russia. Individuals experienced heightened uncertainty and increased their food purchases. In addition, there were instances where individuals spent money to relocate.

In this study, our objective was to address the problem of identifying the dynamics of customer banking behavior, as influenced by macroeconomic factors and seasonal variations. Specifically, our analysis focused on the amount of spending by customers in various currencies as an indicator of their behavior.

2.2 Econophysics and Lagrangian formalism

Economics shares numerous similarities with physics, enabling the application of physical terminology and established methods from physics to solve economic problems. Energy is a fundamental concept in physics that enables bodies to move and accumulate potential energy that can be realized at a later time. For instance, just as a ball held above the ground possesses potential energy, enterprises, factories, and economic actors have the capacity to accumulate resources that can later convert into income or losses. This implies that when the ball is released, its potential energy transforms into kinetic energy, until the system reaches a state of minimum energy.

However, the question arises regarding the selection of a unit of energy measurement in the economy [18]. For instance, in work [10], money serves as the unit of measurement, representing currency. Consequently, the energy exchange

ICCS Camera Ready Version 2025
To cite this paper please use the final published version:
DOI: 10.1007/978-3-031-97626-1_17

is interpreted as a commodity-money relationship between two economic agents. Without external energy, a physical system that has attained its lowest possible energy state cannot transition to a state of higher energy.

This concept corresponds to an investment in the company's economy. When investments are low or loan rates are high, the company is unable to develop. Consequently, the only resource it can utilize is its accumulated capital, which can be likened to potential energy.

In our work, we also considered money as energy, but we adopted an abstract concept as units of measurement, such as γ. Different tasks may require varying values of γ. It can be represented as money, quantity, and other economic units [18]. We assumed that γ serves as an arbitrary unit to obtain the appropriate coefficients in equations.

We considered the time series of the length T, which corresponds to the amount of transactions accumulated by all clients. The X defines our system, i.e., there is one degree of freedom, where X is the generalized coordinate. The unit of measurement for the amount X is an arbitrary currency β (in our cases rubles). In fact, we included all the transformations we performed on the original time series in this variable. Specifically, during training the neural network, we normalized the values to the range from -1 to 1.

The generalized velocity $\dot{X}$ represents the rate of change in the amount of transactions over one period Ω, with units of $\frac{\beta}{\Omega}$. The duration of change can be equivalent to a day, a week, or any other number of days, depending on the trend smoothing period employed.

The kinetic energy is the quadratic form of the generalized coordinates, that is, in the one-dimensional case is equal to $T = \frac{1}{2}m\dot{X}^2$. The inertial mass of manufacturing encompasses all elements that suppress alterations in the manufacturing process, including inflexible technology and difficulties in adjusting manufacturing inputs [10]. This mass should have units of measurement such that the quantity $m\dot{X}^2$ is measured in γ, that is, $\frac{\gamma\Omega^2}{\beta^2}$.

One possible method to derive the dynamic equation is to employ the Lagrangian formalism. For most physical systems, it is feasible to determine its Lagrangian. The Lagrangian is a function that encapsulates the behavior of all components within the system. Once the Lagrangian is known, the dynamic equation can be derived through the Euler-Lagrange equation with dissipative function F,

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{X}} = \frac{\partial L}{\partial X} - \frac{\partial F}{\partial \dot{X}}. \tag{1}$$

The primary challenge in constructing the model lies in determining the suitable potential energy function to derive the dynamic equation. To address this, we analyzed the system's behavior through the Taken's embeddings in the subsequent section.

2.3 Embeddings and phase trajectories

Let us assume that the dynamics of the system is governed by a multidimensional system under an ordinary differential equation $\dot{y} = v(x)$, where $x \in \mathcal{M}$ is the

ICCS Camera Ready Version 2025
To cite this paper please use the final published version:
DOI: 10.1007/978-3-031-97626-1_17

state, $\mathcal{M}$ is the n-dimensional smooth compact manifold and v is the vector field defined on $\mathcal{M}$, i.e., v is an element of the tangent bundle $T\mathcal{M}$. Under the initial condition $z \in \mathcal{M}$, the solution of the system is a flow $g^t : \mathcal{M} \to \mathcal{M}$. This map represents the one-parametric group of diffeomorphisms $\{g^t\}_{t\geq 0}$ such that $v(x) = \frac{d}{dt}|_{t=0}(g^t x)$.

We denoted the projections of this system as $\{\alpha(g^t)\}_{t\geq 0}$, where $\alpha : \mathcal{M} \to \mathbb{R}$ is the projection function which is so-called an observation function. In our case, these projections represent the volume of customer expenses. Our objective was to analyze the values of the observation function and understand the qualitative behavior of the system. This objective realized through the application of Taken's theorem.

We refered to the embedding as the map, which is defined as follows: the differentiable map $f : \mathcal{M} \to \mathcal{N}$ is called immersion if $d_x f : T_x\mathcal{M} \to T_{f(x)}\mathcal{N}$ is an injective map for all $x \in \mathcal{M}$; then embedding is the immersion which is also the homeomorphisms between $\mathcal{M}$ and $f(\mathcal{M}) \subset \mathcal{N}$. We denoted time delay as τ.

Theorem 1 (Taken's theorem). *Suppose $\tau \geq 0$ and $m \geq 2n+1$. Let g^t satisfy the conditions:*

(i) There is no periodic orbits of the period τ and 2τ.
(ii) There is a finite number of periodic orbits of the period $k\tau, k \in \mathbb{N}, k \geq 3$.
(iii) If $v(x) = 0$, then eigenvalues of $d_x g^t$ are all distinct and not equal to one.

Then $\mathcal{F}_{\alpha,g,\tau} : \mathcal{M} \to \mathbb{R}^m$, such that $\mathcal{F}_{\alpha,g,\tau}(x) := (\alpha(x), \alpha(g^\tau x), \ldots, \alpha(g^{(m-1)\tau} x)) \in \mathbb{R}^m$ is an embedding.

Thus, by establishing a specific phase flow, it is possible to construct an embedding that is devoid of self-intersections and uniquely corresponds to the initial dynamics of the system. Utilizing this embedding, we interpreted the behavior of a system that we do not directly observed, provided we restricted our observations to specific projections.

If the time series X has the one-dimensional observations and T data points, then the embedding's values are equal to

$$\begin{pmatrix} X(t) & X(t+\tau) & \ldots & X(t+(m-1)\tau) \\ X(t+1) & X(t+\tau+1) & \ldots & X(t+(m-1)\tau+1) \\ \vdots & \vdots & \ddots & \vdots \\ X(t+(m-1)) & X(t+\tau+(m-1)) & \ldots & X(t+(m-1)\tau+(m-1)) \end{pmatrix}, \quad (2)$$

where m is the embedding dimension, τ is the time delay and $t = 1, \ldots, T - (m-1)\tau - (m-1)$.

We assumed the time delay $\tau = 1$. In our tasks, initially the data is already discrete. We have selected days for transaction date measurements by grouping the hours and minutes during which processing of the source data takes place, and then smoothing the results. In other words, we have chosen this specific sampling period based on the requirements of the task at hand. We determined the optimal embedding dimension using false nearest neighbors [19].

ICCS Camera Ready Version 2025
To cite this paper please use the final published version:
DOI: 10.1007/978-3-031-97626-1_17

Since the embedding dimension does not display the entire trajectory of the system, we employed the dimensionality reduction method, t-distributed stochastic neighbor embedding [20] (t-SNE), to display the phase trajectory of the system on a plane.

Analyzing Figure 2, we noted the cyclic pattern in the transactions trajectory. The colors in the graph correspond to the years, commencing with 2018 (brown) and concluding with 2022 (blue). The graph's points align with the annual cycles, as evident in instances such as January 2020, 2021, and 2022. The Figure shows a deviation in the period for the month of April in 2020 due to the COVID-19 pandemic.

It is also noteworthy that the cycles of March in 2018, 2019, 2020, and 2021 exhibit similar patterns of behavior, with points exhibiting these patterns having close orbits. In contrast, the month of March 2022 stands out due to the outbreak of the Russian-Ukraine conflict. The trajectory's movements are directed towards the point marked for April 1, 2022. During these months, the system encounters external influences and deviates from its typical phase trajectories.

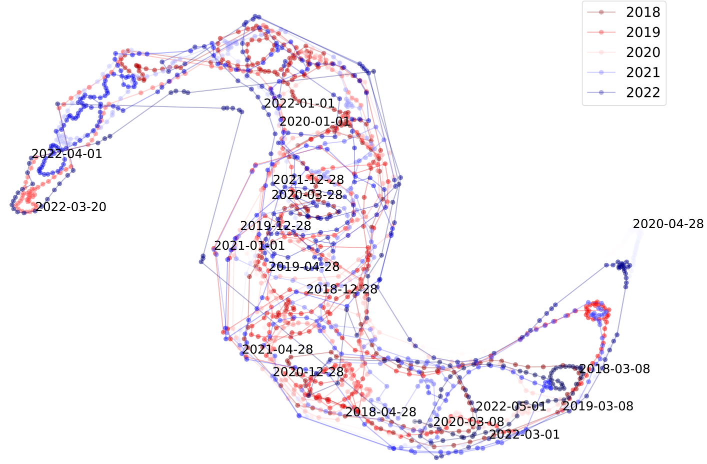


Fig. 2: Projection of the phase portrait of a system onto a plane using the dimensionality reduction method t-SNE.

This structure of orbits resembles periodic fluctuations. One of the most basic models in physics, which describes periodic oscillations, is the mathematical pendulum model. However, in addition to the seasonal fluctuations, there are also some irregularities that occur in the crisis state of the economy. Therefore, we have taken into account a pendulum that is affected by an external field, which varies over time. This external field can be understood in different ways, namely store closures, tax increases, restrictions on public gatherings, negative news events, and so forth. Finally, it also makes sense to assume a damping effect

ICCS Camera Ready Version 2025
To cite this paper please use the final published version:
DOI: 10.1007/978-3-031-97626-1_17

in our system, since if we do not affect the system at all, it will go into some kind of equilibrium economic state, which we do not observe in the real world due to a large number of random factors and geopolitical and macroeconomic decisions.

2.4 Dynamic equation

In the absence of external forces and frictional forces, the pendulum will persist in a specific oscillatory pattern. However, under the persistent influence of frictional forces, the pendulum will gradually decline and reach a state of equilibrium. Subsequently, upon the introduction of even a minor external perturbation, the pendulum will deviate and, due to the frictional forces, gradually decelerate until it returns to its usual state and, thereby, achieves a state of equilibrium.

In our interpretation, the pendulum is perpetually affected by various factors. Crisis situations correspond to significant disruptions, which induce the pendulum to adopt a novel oscillatory mode. However, due to the inherent forces of friction, it inevitably returns to a stable equilibrium, where fluctuations are solely sustained by seasonal perturbations.

Thus, it is feasible to analyze the potential energy in a manner analogous to a pendulum.

$$U(t, X, \dot{X}) = \frac{kX^2}{2} - F(\dot{X}) - \psi(t)X, \tag{3}$$

where $F(\dot{X}) = \frac{\alpha \dot{X}^2}{2}$ is related to a dissipate function, $\psi(t)$ represents the external forces (non-homogeneous function in the right hand side of the dynamic equation), k is the «spring» constant (the measure of the stiffness of the system).

In our case, the «spring's» stiffness can be understood as the extent with which bank customers' transaction patterns return to their pre-shocked state after experiencing a shock. The higher the «spring» coefficient, the faster the customers return to their original spending levels.

After analyzing the seasonal patterns of X, we identified the two seasonal components with periods of 12 and 31. The first component corresponds to monthly fluctuations, while the second component corresponds to daily fluctuations. We decided to model these components through the periodic functions $\cos(\frac{2\pi}{12}i)$ and $\cos(\frac{2\pi}{31}j)$, where $i \in \{1, \ldots, 12\}, j \in \{1, \ldots, 31\}$. The amplitudes of the oscillation are equals to c_m and c_d, correspondingly. Given the continuous nature of time, we recalculated the frequencies in the cosine argument to ensure that the argument remains a function of time continuously, where $t \in [0, 1]$. We denoted new frequencies as ω_d and ω_m for day and month periods.

Finally, the potential energy is equal to

$$U(t, X, \dot{X}) = \frac{kX^2}{2} - \Big[\phi(t) + c_n f(t)\Big]X - \frac{\alpha \dot{X}^2}{2}, \tag{4}$$

where $\phi(t) = c_d \cos(\omega_d t) + c_m \cos(\omega_m t)$.

ICCS Camera Ready Version 2025
To cite this paper please use the final published version:
DOI: 10.1007/978-3-031-97626-1_17

The Lagrangian of the system equals to

$$L = \frac{1}{2}m\dot{X}^2 + \psi(t)X - \frac{kX^2}{2} - \frac{\alpha\dot{X}^2}{2}. \tag{5}$$

After taking derivatives in (5) and substituting them to (1), the resulting dynamic equation is

$$\ddot{X} = -\frac{k}{m}X + \frac{1}{m}\psi(t) - \frac{\alpha}{m}\dot{X}, \tag{6}$$

or if we redefine the terms as $\frac{k}{m} = b$, $\frac{\alpha}{m} = a$,

$$\ddot{X} + a\dot{X} + bX = \frac{1}{m}\psi(t). \tag{7}$$

Given that we calibrated the parameters a, b and determined the function ψ such that the solution of the dynamics equation best aligns with the initial data, the term m disregarded, as it incorporated into the function ψ. Alternatively, it is possible to set m to 1 by using a re-normalized units of measurement.

3 Results

We evaluated the quality of our model with two steps. Firstly, we assessed whether the solution of the dynamic equation with non-homogeneity, restored through the PINN, aligns with the initial data. Secondly, we investigated how the non-homogeneity function could be predicted using the news background data. For the network architecture we used the inverse physics-informed neural network with three feed forward layers and tanh activation, and with three losses: (i) data loss; (ii) initial conditions; (iii) equation loss to satisfy (7).

We utilized transaction data from debit cards of commercial bank clients in Russia for the period of January 2018 to July 2022. This data encompasses information about approximately $10,000$ active consumers' expenses, which were sampled daily. For news parsing, we utilized a regional news aggregator (`https://sanktpeterburg.bezformata.com`).

3.1 Empirical non-homogeneity function

By employing the appropriate external force signals and periodic components, we discerned the corresponding patterns of customers behavior. Figure 3 shows the periodic signal $\phi(t)$ and shocks $f(t)$ from the economy. The periodic component is achieved by combining two cosine functions with the periods of 12 and 31, which correspond to the year and month seasons. Using the Fourier analysis we also detected seasons with the periods of 7 and 14 days. However, fluctuations with this periodicity do not significantly affect the trend component of transaction amount. Therefore, to accelerate the calculations, they were not included in the analysis. The amplitudes c_d, c_m of cosine functions were determined from the data with the Fourier analysis.

ICCS Camera Ready Version 2025
To cite this paper please use the final published version:
DOI: 10.1007/978-3-031-97626-1_17

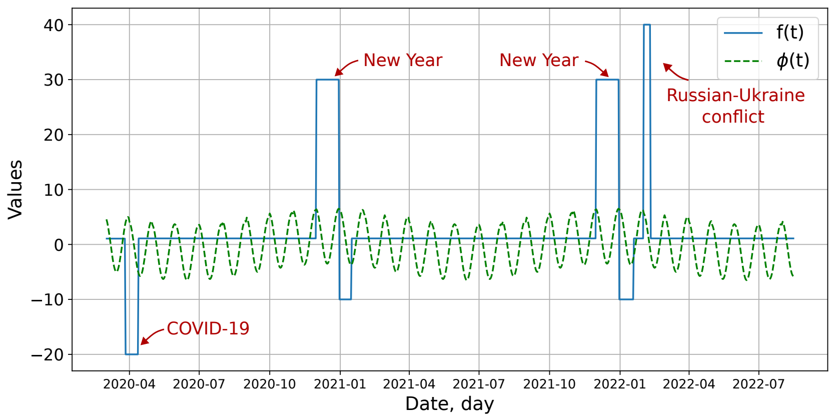


Fig. 3: Non-homogeneity function of the equation with periodic and empirical components. The absolute values on the Y-axis represent the relative influence of external factors. The higher the absolute value, the greater the influence.

Based on empirical considerations, we identified the function $f(t)$. Firstly, during the initial phase of the COVID-19 pandemic (first six months of 2020), purchasing power experienced a substantial decline due to the restrictions on offline shopping and the extended quarantine period. Secondly, during the New Year period, individuals exhibited a significant increase in spending, aligning with the customary holiday season in December. However, after the New Year, the frequency of discounts decreased sharply, prompting individuals to adopt a saving mindset following their substantial spending. Therefore, spending patterns experienced a decline. Lastly, the period commencing in 2022 coincided with the onset of the Russian-Ukraine conflict, during which people actively engaged in purchasing essential commodities.

The numerical solution of the equation (7) with this empirical signals is demonstrated in the next subsection.

3.2 Solution of the dynamic equation

Figure 4 shows the equation solution with empirical signals (red color) obtained using the Runge-Kutta numerical scheme of order 4, related to calibrated parameters determined through the L-BFGS-B optimization method (limited and bound constrained quasi-Newton method of Broyden, Fletcher, Goldfarb, and Shanno [21]). The numerical solution closely approximated the normalized value of expenditures. The solution captured all the significant peaks from the economic shocks, as well as similar seasonal variations.

However, it is significantly more advantageous to obtain these signals directly from the data and endeavor to comprehend the factors that generate such signals. The green color in Figure 4 denotes the PINN's solution, which is the approximation obtained using a neural network. It closely resembles the initial

ICCS Camera Ready Version 2025
To cite this paper please use the final published version:
DOI: 10.1007/978-3-031-97626-1_17

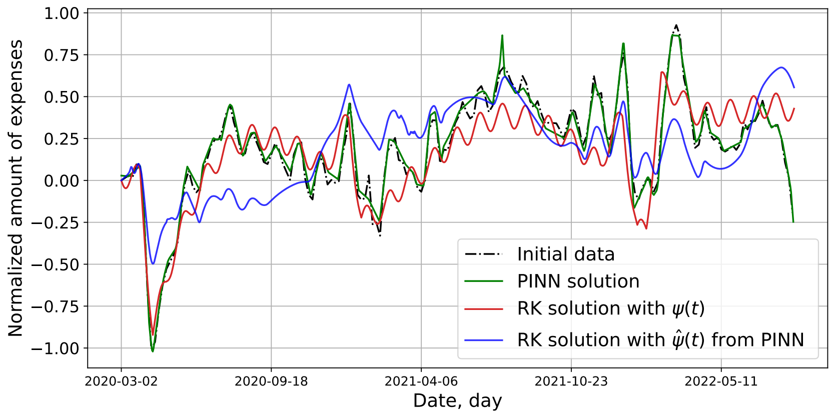


Fig. 4: Solutions obtained using the Runge-Kutta method for the empirical non-homogeneity function $\psi(t)$ and for the function obtained from PINN $\hat{\psi}(y)$.

data (black color) because we solved the inverse task for PINN. However, the most interesting aspect is the non-homogeneity $\hat{\psi}$ achieved during the PINN's training process.

The solution obtained the Runge-Kutta numerical scheme with $\hat{\psi}$ is shown in blue. It is evident that the calibrated function precisely aligns with the dynamic equation, as the solution exhibits all the primary characteristics of the initial data. However, it contains less information regarding the monthly seasonal patterns, capturing more general trends. However, when considering customer behavior in reaction to economic shocks, we wish to capture general trends rather than specific local patterns.

From a practical standpoint, it is crucial not only to identify the non-homogeneity (external impact on the system) but also to explain the underlying causes. Without making purely empirical assumptions about the initial data, it is challenging to ascertain when and how the impact was applied to the system. Therefore, we proceeded to interpret this external influence in further detail.

3.3 The non-homogeneity interpretation by news

To comprehend the impact of news on customers behavior, it is crucial to identify the relevant signals from news sources. To achieve this, we employed data collection techniques from social media and reputable news agencies. Utilizing specialized filters, we identified significant events related to COVID-19 within the specified time frame, spanning from January 14, 2020, to June 30, 2021.

Firstly, we cleaned the news data from non-semantic words using regular expressions. Then, we selected keywords relevant to COVID-19 events, such as virus, vaccine, COVID, epidemic, sick, disease and others. Subsequently, for each day, we employed these keywords to identify sentences related to our topic. Afterward, we counted the number of sentences with the semantic meaning to

ICCS Camera Ready Version 2025
To cite this paper please use the final published version:
DOI: 10.1007/978-3-031-97626-1_17

our topic and divided it by the total number of sentences in our news database for the current date.

Figure 7 presents the computed average number of news articles related to COVID-19. It is noteworthy that news activity experienced an uptick in March 2020, followed by a decline in the summer of the same year, and further there were fluctuations around the value 0.1.

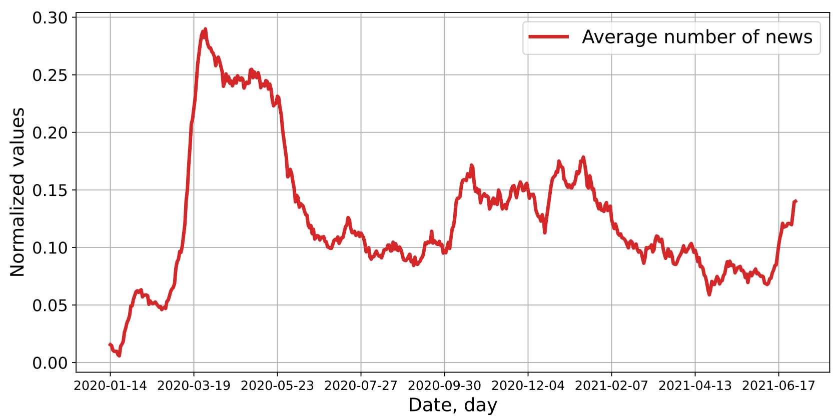


Fig. 5: The average quantity of news coverage over a specified time frame, spanning from January 14, 2020, to June 30, 2021.

To explain the non-homogeneous function within the dynamic equation, a regression model was employed, wherein the independent variables were the dates, expressed in terms of days, months, and years, as well as the mean value of news pertaining to a given topic, derived from the analysis of news datasets.

We employed the gradient boosting regression model from XGBoost library with a total of 100 trees and a maximum depth of 3. To ensure statistical significance, we randomly partitioned the data into training and testing sets with a ratio of $8:2$. This process is repeated 200 times.

To assess the accuracy of the attributes in describing the non-homogeneity function, we utilized the coefficient of determination. This coefficient indicates the proportion of variance that is accounted for in the model's predictions. In our case its average values equals 0.8 with standard deviation 0.09 on the repeated train and test samples. This suggests a good descriptive capacity of the news, given that we have selected them based on fixed regular expressions.

Figure 6 shows the SHAP values. We used four indicators, and as can be seen, the most significant indicator was related to news. A large number of average news has a negative impact on the predicted external effect. The year can also be clearly divided into two parts. However, the indicators associated with the day and month of news have both positive and negative impacts, with some values being large and others small.

ICCS Camera Ready Version 2025
To cite this paper please use the final published version:
DOI: 10.1007/978-3-031-97626-1_17

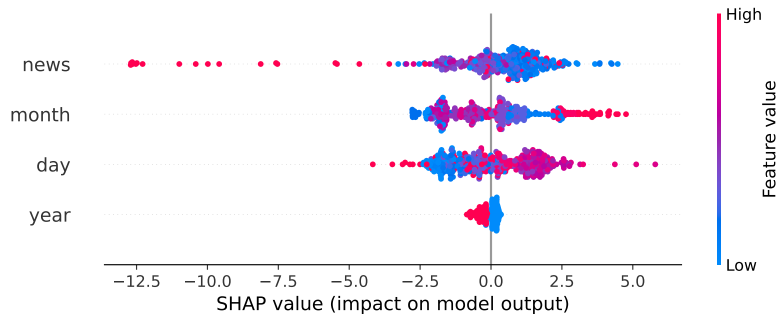
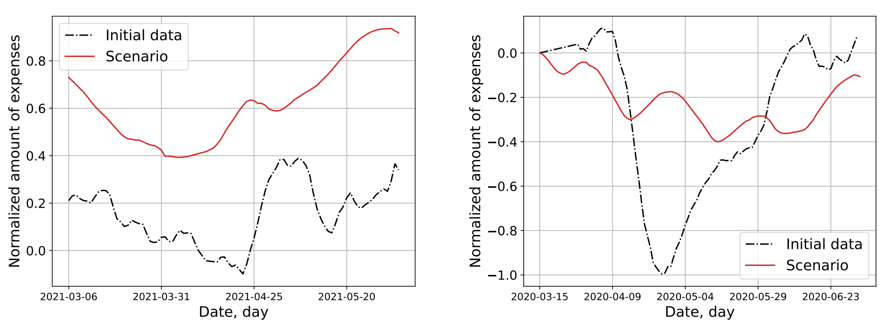


Fig. 6: Feature impact visualization. The feature names are plotted along the Y-axis. The X-axis indicates the impact on the model's prediction. Values above zero indicate a positive impact, while negative values indicate a negative impact. The color coding represents the magnitude of the feature's values.

Figure 7 illustrates the scenarios that were generated by varying the news factor. Specifically, in the left-hand figure, we increased the coverage of news related to COVID-19 by 15% for the period spanning March to June 2021. Then, we predicted a non-homogeneous function and solved the differential equation (7) numerically. As can be seen, the number of expenditures increased during this time period. In another scenario, we reduced news coverage of COVID-19 by 20%. The period selected corresponds to the onset of the pandemic. It can be observed that if news coverage is reduced, people would be less exposed to external influences, and therefore their spending would decrease less. At the same time, seasonal factors remain.



(a) Increased by 15%. (b) Decreased by 20%.

Fig. 7: Scenario modeling based on the variation of the amount of news about COVID-19.

ICCS Camera Ready Version 2025
To cite this paper please use the final published version:
DOI: 10.1007/978-3-031-97626-1_17

4 Discussion and Conclusion

In this study, we have addressed the issue of understanding the behavior of bank customers based on transactional data. Rather than relying solely on machine learning algorithms, we employed a hybrid approach that combines physics-informed neural networks and econophysics techniques. A key point in our work is the connection of the economic system through the concepts of Lagrangian mechanics. Analogies with physics have made it possible to analyze the behavior of customers and interpret the external factors with selected set of variables, namely news, day of the week, month, and year. As a result, news proved to be a significant factor, accurately reflecting the state of the national economy.

We analyzed phase trajectories of a system that describes the complete dynamics of human behavior, considering the projections of this system as customer expenditures and applying the Takens' embedding theorem.

At the same time, we believe that further improvements to the description of non-homogeneity can be achieved by considering not only the news context, but also macroeconomic indicators of the country in question. Additionally, it would be interesting to explore the use of LLM-based semantic filtering of news.

Moreover, we conducted scenario simulations, demonstrating how manipulating news inputs can influence consumer behavior. This finding is attributed to individuals' tendency to adapt their spending patterns in response to news events, aiming to optimize their personal utility. We believe that our research can be beneficial for managers at banks and other financial institutions to analyze and forecast the behavior of their customers by modeling various scenarios, including news reports.

References

1. Paramesha, M., Rane, N.L., Rane, J.: Artificial intelligence, machine learning, deep learning, and blockchain in financial and banking services: A comprehensive review. Partners Universal Multidisciplinary Research Journal 1(2), 51–67 (2024)
2. Xu, L., Roy, A.: Cashback as cash forward: The serial mediating effect of time/effort and money savings. Journal of Business Research 149, 30–37 (2022)
3. Vullam, N., Vellela, S.S., Reddy, V., Rao, M.V., SK, K.B., Roja, D.: Multi-agent personalized recommendation system in e-commerce based on user. In: 2023 2nd International Conference on Applied Artificial Intelligence and Computing (ICAAIC). pp. 1194–1199. IEEE (2023)
4. Al-Hashedi, K.G., Magalingam, P.: Financial fraud detection applying data mining techniques: A comprehensive review from 2009 to 2019. Computer Science Review 40, 100402 (2021)
5. Zakharov, K., Stavinova, E.: Time-dependent differential privacy for enhanced data protection in synthetic transaction generation. In: Proceedings of the 2024 13th International Conference on Software and Computer Applications. pp. 112–117 (2024)
6. Wuisan, D.S., Handra, T.: Maximizing online marketing strategy with digital advertising. Startupreneur Business Digital (SABDA Journal) 2(1), 22–30 (2023)

ICCS Camera Ready Version 2025
To cite this paper please use the final published version:
DOI: 10.1007/978-3-031-97626-1_17

7. Khandani, A.E., Kim, A.J., Lo, A.W.: Consumer credit-risk models via machine-learning algorithms. Journal of Banking & Finance 34(11), 2767–2787 (2010)
8. Saretzki, J., Pretsch, J., Pretsch, E., Grossmann, G.: Experience, behavior, and action in crisis situations: A literature review. European Journal of Social Sciences Education and Research Articles 8 (2021)
9. Koshkareva, M., Kovantsev, A.: Crisis behaviour strategy recognition using transactional data. Procedia Computer Science 229, 208–217 (2023)
10. Estola, M., Dannenberg, A.A.: Newtonian and lagrangian mechanics of a production system. Hyperion International Journal of Econophysics New Economy 9(2), 7–26 (2016)
11. Sandler, U.: S-lagrangian dynamics of many-body systems and behavior of social groups: Dominance and hierarchy formation. Physica A: Statistical Mechanics and its Applications 486, 218–241 (2017)
12. Zakharov, K., Stavinova, E., Lysenko, A.: Trgan: A time-dependent generative adversarial network for synthetic transactional data generation. In: Proceedings of the 2023 7th International Conference on Software and e-Business. pp. 1–8 (2023)
13. Kumari, L., Bhattacharjee, K., Sharma, N., Kumar, S., Kumari, A.: Machine learning models in customer behaviour prediction: A comparative analysis. In: 2024 7th International Conference on Contemporary Computing and Informatics (IC3I). vol. 7, pp. 957–959. IEEE (2024)
14. Harish, V., Benitta, D.A.: Construction of prediction model and forecasting trends based on consumer behaviour. In: 2024 OPJU International Technology Conference (OTCON) on Smart Computing for Innovation and Advancement in Industry 4.0. pp. 1–6. IEEE (2024)
15. Deniz, E., Bülbül, S.Ç.: Predicting customer purchase behavior using machine learning models. Information Technology in Economics and Business 1(1), 1–6 (2024)
16. Sakthi, B., Sundar, D.: An effective customer behavior prediction system in digital marketing using ensemble serial cascaded network. In: 2024 International Conference on Intelligent Algorithms for Computational Intelligence Systems (IACIS). pp. 1–8. IEEE (2024)
17. Guleva, V.Y., Kovantsev, A.N., Chunaev, P.V., Gornova, G.V., et al.: Value-based modeling of economic decision making in conditions of unsteady environment. Journal Scientific and Technical Of Information Technologies, Mechanics and Optics 147(1), 121 (2023)
18. Zakharov, K., Kovantsev, A., Boukhanovsky, A.: Coupling of lagrangian mechanics and physics-informed neural networks for the identification of migration dynamics. Smart Cities 8(2), 42 (2025)
19. Kennel, M.B., Brown, R., Abarbanel, H.D.: Determining embedding dimension for phase-space reconstruction using a geometrical construction. Physical review A 45(6), 3403 (1992)
20. Van der Maaten, L., Hinton, G.: Visualizing data using t-sne. Journal of machine learning research 9(11) (2008)
21. Byrd, R.H., Lu, P., Nocedal, J., Zhu, C.: A limited memory algorithm for bound constrained optimization. SIAM Journal on scientific computing 16(5), 1190–1208 (1995)

ICCS Camera Ready Version 2025
To cite this paper please use the final published version:
DOI: 10.1007/978-3-031-97626-1_17