# References to Academic Work

# THIS IS A CONFIDENTIAL FILE THAT CAN ONLY BE USED BY THE ADMISSION COMMITTEE

Author: Boxuan Huang (bh2821@nyu.edu)

**What Are Included:**

Inside this document, a part of public and non-public academic achievements, research achievements, and undergraduate project reports of this applicant are presented in a translation form (since the original copies are all in Chinese, they were all translated by the applicant himself). While there is no guarantee that every detail presented will be identical to his original version, the revealed main ideas, illustrated bullet points, and documented technical details have all been double-checked without errors to have adequate evaluation values. Inside the body part of this document, 4 selected paragraphs from academic texts are created solely by the applicant himself (while some unprovided parts might have been a collaborative result of a certain team), with 4 published or unpublished academic results included and 4 more unpublished undergraduate project or teamwork reports that can be additionally requested from the applicant. All of those academic texts are highly associated with the academic context of the applied program, or materials that can robustly demonstrate the applicant's abilities that are essential, or beneficial to the postgraduate study, and it should be reiterated that any content from this document can and can only be used with an evaluation aim by the particular admission committee that was distributed to.

A brief index of academic texts included is presented below.

**Index:**

**Academic and Research Result (included passages):**



**Undergraduate Project or Reports (you have to request):**



It is kindly reminded that the academic samples are not limited to those listed above. Therefore, if your admission committee is interested in seeing more of some academic details within a particular field, or conducting a conversation to learn more about the content of the selected passages, please feel free to contact me via email.

# 1. [A Published Patent] A Hybrid Encryption-decryption System and Methods of Multi-step Layered Encryption and Multi-dimensional Distribution

**Description**------------------------------------------------------------------------------------------------------------------------

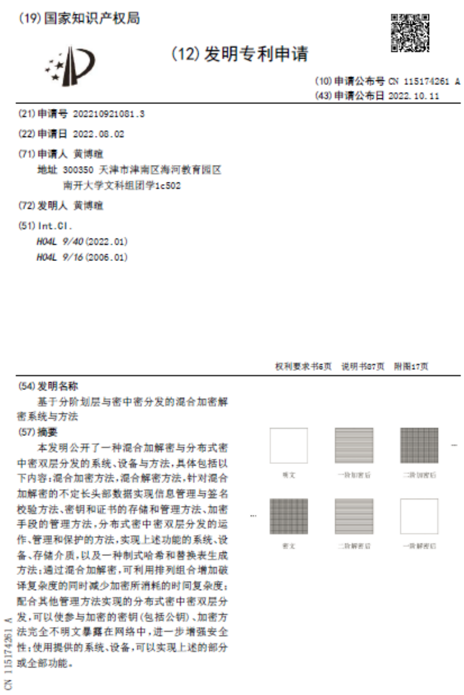
Patent Number: CN202210921081

Patent Publication Number: CN115174261 A

The following parts are selected and translated from the public patent content.

Please note that a simple implementation has been developed and open-sourced on GitHub where anyone can freely get a copy of this encryption program, with an Apache-2.0 license that simultaneously protects my patent rights.

The present invention discloses a system, device, and method for a multi-step layered encryption and multi-dimensional distribution which specifically includes the following contents: multi-step layered encryption method, multi-step layered decryption method, information management and signature verification method for variable length header data of multi-step layered encryption and decryption, storage, and management method for keys and certificates, management method for encryption means, operation, management, and protection method for distributed multi-dimensional distribution of encryption and decryption, a system, device, storage medium for implementing the above functions, and a method for generating standard hashes and replacement tables. With such hybrid encryption and decryption, permutation and combination can increase the decoding complexity while reducing the time complexity consumed by encryption. The multi-dimensional distribution of encryption in combination with other management methods can expose the encryption keys (including public keys) and encryption methods completely in the network, further enhancing security. By using the systems and equipment provided, some or all of the above functions can be achieved.


(19) 国家知识产权局

(12) 发明专利申请

(10) 申请公布号 CN 115174261 A
(43) 申请公布日 2022.10.11

(21) 申请号 202210921081.3
(22) 申请日 2022.08.02
(71) 申请人 黄博暄
地址 300350 天津市津南区海河教育园区 南开大学文科组团学1c502
(72) 发明人 黄博暄
(51) Int.Cl.
H04L 9/40 (2022.01)
H04L 9/16 (2006.01)

权利要求书6页 说明书87页 附图17页

(54) 发明名称

基于分阶划层与密中密分发的混合加密解密系统与方法

(57) 摘要

本发明公开了一种混合加解密与分布式密中密双层分发的系统、设备与方法，具体包括以下内容：混合加密方法，混合解密方法，针对混合加解密的不定长头部数据实现信息管理与签名校验方法、密钥和证书的存储和管理方法、加密手段的管理方法，分布式密中密双层分发的运作、管理和保护的方法，实现上述功能的系统、设备、存储介质，以及一种制式哈希和替换表生成方法；通过混合加解密，可利用排列组合增加破译复杂度的同时减少加密所消耗的时间复杂度；配合其他管理方法实现的分布式密中密双层分发，可以使参与加密的密钥(包括公钥)、加密方法完全不明文暴露在网络中，进一步增强安全性；使用提供的系统、设备，可以实现上述的部分或全部功能。

CN 115174261 A

**Literature Selection**----------------------------------------------------------------------------------------------------------

**Background Technology:**

…

At present, Internet-based information transmission has permeated all aspects of our society. A multitude of industries with high security and confidentiality requirements, from digital finance to information technology and national defense, are increasingly relying on the Internet as a medium for information acquisition, storage, and exchange, hence ensuring the security of Internet information has become a critical issue in the Internet strategy of organizations, businesses, and even the whole nation. Unfortunately, the development of information security on the Internet now faces a limitation due to a variety of factors. The rapid growth of modern computer computing power is increasingly challenging the security of traditional cryptographic protection methods. On the other hand, while quantum computers and quantum encryption technologies offer promising prospects for protecting information security, the process of achieving the transition to quantized protection in the industry is undeniably not overnight, regardless of how long it may take from initial development to maturity.

With the principle of reducing "menu costs", improving existing technology to achieve higher security protection has become a significant direction in cryptography research in recent years, and among them, the application of public-key cryptography, a technology that allows two parties to encrypt and decrypt information via asymmetric keys, with the private key not being transmitted over the Internet, has become an epoch-making advance in modern Internet security. As an alternative solution, in order to reduce the time consumption of encryption and decryption, public-key cryptography is universally used to encrypt the session key of symmetric cryptography, and then the session key of symmetric cryptography can then be used to encrypt the real message, which achieves a double trade-off between security and speed. In order to overcome the approach mentioned above, multiple encryption methods are gradually being generalized and applied, for example, typical applications such as 3DES and nested-salted hashing illustrate the promise of such improvements, which hopes that multiple encryptions using different keys can enable permutation-based functions simultaneously to increase the complexity of the process of encryption.

However, since the above three known improvement methods have been mature for many years, their strategies have been thoroughly studied by the industry and their flaws have been increasingly exposed and demonstrated. For example, the fatal problem in public-key cryptography is the disclosure of the public key. In many cases, the inevitability of public and private key pairs leads to insufficient security of ciphertexts, while it is theoretically possible to obtain plaintexts by loop operations when both ciphertext and public key are available. Another public key application, which encrypts the session key for symmetric cryptography, is on the other hand also imperfect. In certain circumstances, the deciphering of a password can skip the steps of public key cryptography encryption and directly crack the message used for symmetric encryption - in this case, the security of the data directly depends on the security performance of the symmetric encryption method, and under brute force cracking, the improvement of security is almost independent of the participation of public key cryptography. Although the idea of using multiple encryption methods to exchange spatial and temporal complexity for higher security due to permutation and combination is desirable, traditional methods that use the same algorithm and encrypt multiple times at a fixed number of times still have problems such as fixed rules and insufficient security improvement due to excessively simple, fixed permutation and combination strategies.

…

**Summary of the Invention:**

…

To achieve the above objectives, the technical solution adopted by the present invention is:

In the first aspect, the embodiment of this invention provides a hybrid encryption method (H-encryption method) to achieve the use of existing various cryptographic methods, which reduces the complexity of encryption and decryption time as much as possible through hierarchical layers, permutations, and combinations, while bringing a geometrically increasing high decryption complexity, and encrypting plaintext into ciphertext.

This method comprises:

Select a total number of *Step* $m$ and a *Layer* $n_i$ ($1 \leq i \leq m$ and $i \in Z$) for each *Step* $i$, which can be recorded in vector form as follows:

$$n_1 \ n_2 \ \cdots \ n_{m-1} \ n_m;$$

Select an encryption method and key for each *Layer* of each *Step*;

Enter an iteration;

Starting from the first *Step* $i=1$: If the *Step* has only one *Layer,* the block size for this *Step* is equal to the layer size (encryption throughput). The plaintext length is divided by the block size to obtain the number of block partitions, which are rounded if they are not evenly divided. For non-tail complete blocks, the encryption method chosen by that layer is used to encrypt the entire plaintext of the full text. If it cannot be evenly divided when partitioning, it is not until the first layer that needs to be filled is encountered, record the number of layers to be filled (1 for plaintext recording), record the body-code rules or length to be filled, then fill in the content of that layer and use the encryption method selected by that layer to encrypt the content of the one-time encryption throughput length of that layer, then end the encryption of this level and perform $i$ autoincrement ($++i$). If the *Step* has multiple *Layers*, the block size for this *Step* is equal to the sum of all layer sizes (encryption throughput) of this *Step*. The number of block partitions is obtained by dividing the length of plaintext by the size of the block. If it is not an integer division, it is rounded up. For non-tail complete blocks, according to the division of layers within the block, each layer is sequentially encrypted using the selected encryption method of that layer to encrypt the content of the one-time encryption throughput length of that layer, completing the encryption within a certain block, and use the same method in sequence to encrypt the next non-trailing complete block. If there is a situation where the remaining plaintext length is less than the block throughput length, that is when there is a need to fill in encryption, the content of the one-time encryption throughput length of each layer that does not need to be filled in the block should be encrypted using the encryption method selected by that layer according to the division of layers within the block. At this point, if the partition cannot be evenly divided, then when the first layer to be filled is encountered, the number of layers to be filled, the body-code rules or length to be filled, and then the content of that layer is filled and the selected encryption method of that layer is used to encrypt the content of the first encryption throughput length of that layer. Then, the encryption of this level is ended, even if there are other layers after that layer of the partition, skip and perform $i$ autoincrement ($++i$);

When *Step* $i \geq 2$, the encrypted ciphertext of the previous order is used as the plaintext to be encrypted in the current order, repeating the same encryption work as when $i=1$, and performing $i$ autoincrement ($++i$);

Loop until $i=m+1$, jump out of the loop, which means the whole encryption has been completed, and the ciphertext encrypted to order $i=m$ is the final ciphertext output.

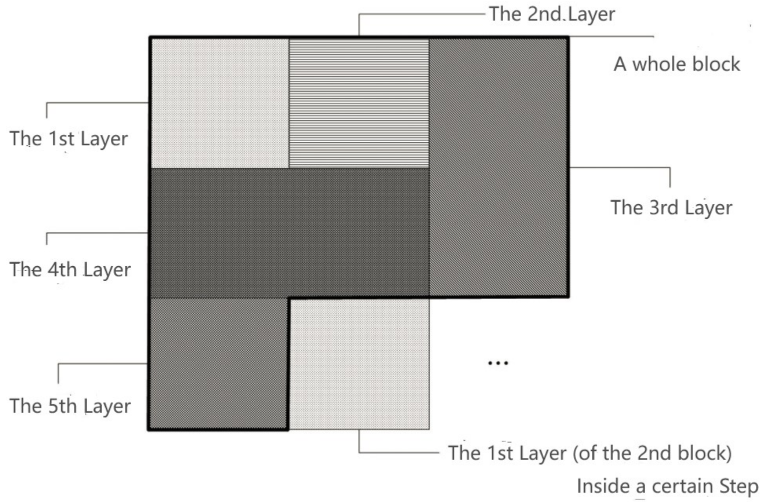
When applied to different scenarios and purposes, there are certain limitations to the encryption methods selected for each level and layer of the transmitted message:

When applied to P2N scenarios, due to the fact that the receiver has multiple parties and the sender has only one party, the encryption methods used for any *Step* or *Layer* of the message to be transmitted are limited to only using any symmetric encryption methods, including symmetric encryption methods for transmitting keys on the network, and symmetric encryption methods for not using network transmission keys. At this point, no asymmetric encryption method is allowed for any *Step* or *Layer* of the transmitted message.

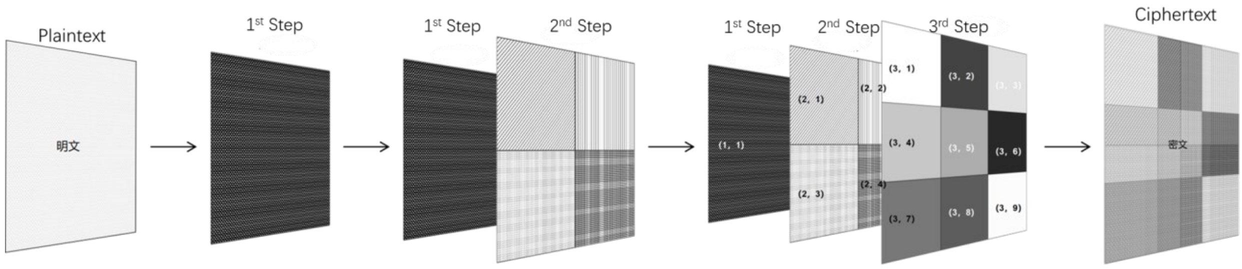
When applied to N2P scenarios, due to the fact that the receiver has only one party and the sender can have multiple parties, the encryption methods used for any *Step* or *Layer* of the transmitted message are limited to only using any symmetric encryption methods, asymmetric encryption method using the receiver's public key, and encryption method using the receiver's public key to encrypt the symmetric key. At this point, asymmetric encryption using the sender's private key encryption or symmetric encryption using key exchange are not allowed for any *Step* or *Layer* of the message to be transmitted.

When applied to P2P scenarios, since there is only one party at the receiver and one party at the sender, the encryption used for any *Step* or *Layer* of the transmitted message is allowed to use any symmetric encryption methods and any asymmetric encryption methods.

For instance, a typical procedure of the application of the 3 *Steps*' encryption, with (1,4,9) *Layers* for each *Step* is illustrated below:




When completed each *Step*, the relationship between the *Steps* is shown in the next graph:




…

**Specific Embodiment:**

…

In order to clearly explain the purpose of the present invention, the internal mechanism, and the process of its implementation, the following will provide a clear and complete description of the technical solution of the implementation of the present invention, together with accompanying drawings. It is important to note that the embodiment described below is only a portion of the embodiment of the present invention, not all of it. In addition, based on the embodiment provided in the present invention, all implementations obtained by ordinary technical men in the arts without the need for creative labor fall within the protection of the present invention.

It should be reiterated that the pseudocode computer language used in the following specific embodiment is the C++ language based on Windows. These uses are due to ease of implementation and expression and do not mean that the protection of the present invention is limited to the C++ language, the Windows operating system, and the following examples. Any embodiment based on the present invention implemented in any computer language, operating system, and form is within the scope of the protection of the present invention.

…

Embodiment One:

According to the provided Figures 1 and 2, the program flow used to solve the proposed hybrid encryption method (H-encryption method) problem is from the first *Step 1* to the outermost *Step m*, each block of land within each *Step*, and each block is encrypted *Layer* by *Layer* from the first *Layer* to the $n_i$ *Layer*. The general form of the specific process is as follows:

…

Transform the process of general form H-encryption into program pseudocode, namely:

```
// ENVIRONMENT PREPARATION ////////////////////////////////////////////////////////////////////////////////////
// Select a Step number m
int STEP = m;
int LAYER [m];
// Select Layer numbers for each Step
LAYER[0]=n1; … ; LAYER[i+1]=ni; LAYER[j+1]=nj; …; LAYER[m-1]=nm;
// Assume that the encryption function for each Step and Layer has a form listed below
void Func_STEP1_LAYER1(…); … ; void Func_STEP1_LAYERn1(…);
…
void Func_STEPm_LAYER1(…); … ; void Func_STEPm_LAYERnm(…);
// Set a key or key pairs for each Step and Layer
unsigned char* Key_STEP1_LAYER1; … ; unsigned char* Key_STEP1_LAYERn1;
…
unsigned char* Key_STEPm_LAYER1; … ; unsigned char* Key_STEPm_LAYERnm;
// Use pointers to manage data according to features
unsigned char* Plain_Start_point; // Set a pointer to the first address of the data to be encrypted
unsigned char* Cipher_Start_point; // Set a pointer to the first address of the encrypted data
unsigned char* Plain_point = Plaintext; // Set a pointer to the current location to be encrypted
unsigned char* Cipher_point = new unsigned char […]; // Pointing to the first address of ciphertext

// ENCRYPTION PROCESS ////////////////////////////////////////////////////////////////////////////////////////////
// Iterate, and encrypt from the 1st Step to the mth Step
for (int i = 1; i <= m; ++i)
{
    // Reset pointers, since for each Step, the plaintext differs
    Plain_Start_point = Plain_point; // Reset the pointer to the first address of the data to be encrypted
    Cipher_Start_point = Cipher_point; // Reset the pointer to the first address of the encrypted data
    // The data that needs to be encrypted now is unsigned char * Plain_ Point;
    // The encrypted ciphertext will be stored in unsigned char * Cipher_ Point;
    UINT64 Length_of_Plain = Get_Length_of_Plain(Plain_point); // Get the length of the plaintext
    UINT64 Number_of_block; // Initialize the number of blocks
    int block_size = 0; // Initialize the size of each block
    // When this particular Step contains only one Layer
    if (LAYER[i-1] == 1)
    {
    … // Truncated, but similar to the provided codes
```

```
}
// When this particular Step contains multiple Layers
else if (LAYER[i-1] > 1)
{
     // The block size is equal to the sum of all layer sizes (encryption throughput)
    {
        block_size += input_size_of_Func_STEPi_LAYER1;
        …
        block_size += input_size_of_Func_STEPi_LAYERn1;
    }
    if((Length_of_Plain/block_size)!=((double)Length_of_Plain/(double)block_size))
        Number_of_block = (Length_of_Plain / block_size) + 1;
    else
        Number_of_block = Length_of_Plain / block_size; // Calculate the number of blocks

    // Officially start encrypting from the 1st block
    int j = 1;
    // If it is a non-tail complete block
    while(j < Number_of_block)
    {
         // Encrypt the 1st Layer inside this block
        {
            Func_STEPi_LAYER1(Plain_point, Cipher_point);
            Plain_point += input_size_of_Func_STEPi_LAYER1;
            Cipher_point += output_size_of_Func_STEPi_LAYER1;
        }
        … // Encrypt other Layers in sequence
        // Encrypt the last Layer inside this block to complete
        {
            Func_STEPi_LAYERni(Plain_point, Cipher_point);
            Plain_point += input_size_of_Func_STEPi_LAYERni;
            Cipher_point += output_size_of_Func_STEPi_LAYERni;
        }
        j++;// move to the next block
    }
    // When jumping out of the loop, assert that we have reached the last incomplete block
    assert(j == Number_of_block)
```

```
        {
            // When the first Layer to be filled is not encountered
            {
                // Encrypt the 1st Layer inside this block
                Func_STEPi_LAYER1(Plain_point, Cipher_point);
                Plain_point += input_size_of_Func_STEPi_LAYERn1;
                Cipher_point += output_size_of_Func_STEPi_LAYERn1;

            … // Encrypt other Layers in sequence
                // Encrypt Layers that do not require padding
                Func_STEPi_LAYERni(Plain_point, Cipher_point);
                Plain_point += input_size_of_Func_STEPi_LAYERni;
                Cipher_point += output_size_of_Func_STEPi_LAYERni;
            }
            // When the first Layer to be filled is encountered
            unsigned char State_Matrix [input_size_of_Func_STEPi_LAYERnj];
            // Create a new state matrix to store the filled data to be encrypted
            Padding_Func_STEPi_LAYERnj(Plain_point, State_Matrix, Padding_Info); // Pad
            Save(LAYERnj, Padding_Info); // Save the padded data
            Func_STEPi_LAYERnj(State_Matrix, Cipher_point);
            // Use the selected encryption method to conduct encryption
        }
    }
    // When it is not the 1st Step
    if (i > 1)
        delete[] Plain_Start_point; // Release process data after previous encryption
    // When it is not the last Step
    if (i < m)
    {
        // The encrypted ciphertext of the current Step serves as the inscription of the next Step
        Plain_point = Cipher_Start_point;
        // Dynamically allocate space for storing new ciphertext
        Cipher_point = new unsigned char […];
    }
}
// Finished
…
```

The hybrid encryption method shown above has a special single-step joint masking effect, which means that plaintext can only be fully decrypted when all keys are fully known. The implementation process and protection mechanism are as follows:

In order to activate the special single-step joint masking effect to enhance the security of the encryption, when performing the presented hybrid encryption, a symmetric encryption method with only one *Layer* can be used for a certain *Step*. The key used can use other generated symmetric keys from all other *Layers* in other *Steps*, the public key part of the asymmetric key, and the key generated through the key exchange process (with P2P mode only). The key for the unique layer of that order can be generated through a unique and determined method (such as by calling a has function) and what should be done is to use this key to encrypt the plaintext with the only *Layer* of the selected *Step*. In the whole process of hybrid encryption, indeed, any single *Step* can act as the mentioned one-*Layered-Step*, but it is usually implemented with the first *Layer* of the first *Step* that acts as the joint masking *Step*.

As for its mechanism of enhancing its security, by implementing the encryption with only one *Layer* in the certain *Step*, particularly with the first *layer* of the first *Step* that acts as the joint masking *Step,* the selected single-*Layer Step* ensures that the decipher cannot partially decipher some keys used to see any part of the whole ciphertext, and only by deciphering all keys of the whole bundle can a decipher achieve his final goal – to see a part, or the whole ciphertext, which significantly increases the decipher complexity by coercively requiring him to decipher more keys. Especially in instances where a large number of symmetric encryption methods are used, decipherers often face multiple uncertain symmetric keys that need to be cracked simultaneously, making it impossible to independently verify the correctness of the symmetric keys used in a certain *Step* or *Layer* without other keys. Only when multiple uncertain symmetric keys are correct at the same time can the decryption goal be finally achieved (but in fact, before finally getting the correct plaintext, nobody knows whether your trying key is correct or not). Therefore, this approach enhances the complexity of cracking by combination and permutation, and when applied to instances that use a large number of symmetric encryption methods, it can further ensure security by increasing the total trying numbers of testing the combined key bundles.

…

It is worth noting that for the above encryption method, performing the reverse operation yields the decryption method as shown in the procedure listed below:

…

## 2. [Unpublished Mathematical Research] An Iterative Model with Discrete Convergence Properties and Its Further Applications

**Description**-----------------------------------------------------------------------------------------------------------------

Status: unpublished research results

Discovery Time: August, 2023

In my research of infinite series with some toolkits inside an unofficially released version of Statisticality, I have discovered an interesting iterative model that has a property of discrete convergence, which means, no matter where you chose a real number in the axis, no matter how big or small, rational or irrational it is, after a large enough number of steps of iteration, it will finally converge to a fixed value that is only defined by the given constant parameters. It is worthy to say that the model has been proved with an expression similar to an MA process, but excitingly, by setting a particular value for each $\theta$ in MA series, it has a spectacular, controllable, convergence property.

Unfortunately, due to its undisclosed status, I am afraid that I might not be able to disclose any further proof details in this document, but instead, some general descriptions and potential fields of application will be revealed in the following part.

The basic iteration form is shown below:

*Iteration:*

$$x = any$$

$$for\left(iter\ in\ 1\text{:}\ \lim_{n\to\infty} n\right)$$

$$\{$$

$$x = (x + j) \times (e/\pi)^k \quad where\ j, k\ are\ constant\ variables$$

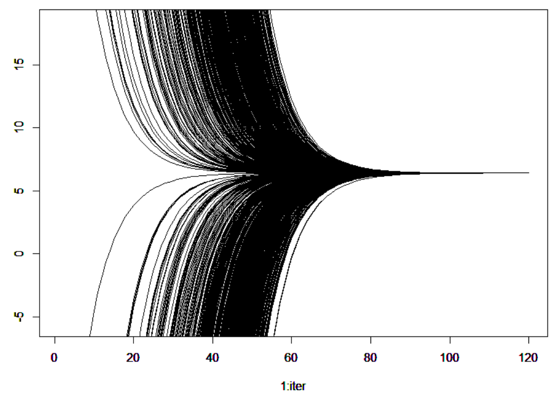
$$\}$$

*It will finally converge to:*

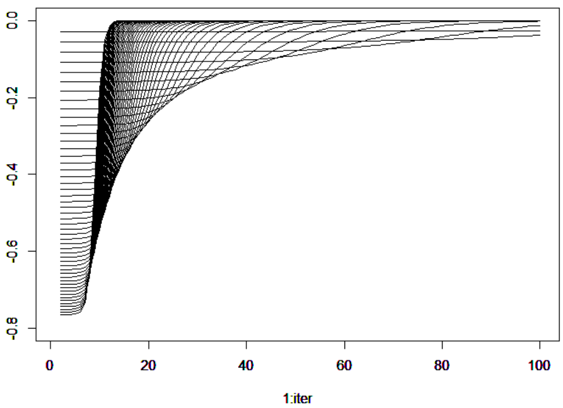
$$x = \frac{j \times \left(\frac{e}{\pi}\right)^k}{1 - \left(\frac{e}{\pi}\right)^k}$$

Do as the procedure illustrated above, there are four steps to perform a complete iteration (if you want to get a whole series, there are additional one step to save the number for each step). First, an arbitrary number of $x$ needs to be initialized, and the selection of $x$ theoretically has no specific requirements, but to be a finite real number. Second, implicitly, we need to choose two constant parameters $j$ and $k$, and normally, the parameter $j$ can be any finite real number, while $k$ should at least be a positive number. But in this case, the choice of its specific value of $j$ or $k$ will accordingly determine the final convergence position of the series. Then, go into the loop and perform the iterative computation for $n$ times, while in theory, times of iteration $n$ should be a number that is large enough, tending to infinity at the limit state. The final step is to get the returned number evaluated as the last iteration of the iterative formula, noted as $x$, converged to a finite real number of *(j*(e/pi)$^k$)/(1-(e/pi)$^k$)*.

**Properties and Potential Fields of Application**----------------------------------------------------------------------------

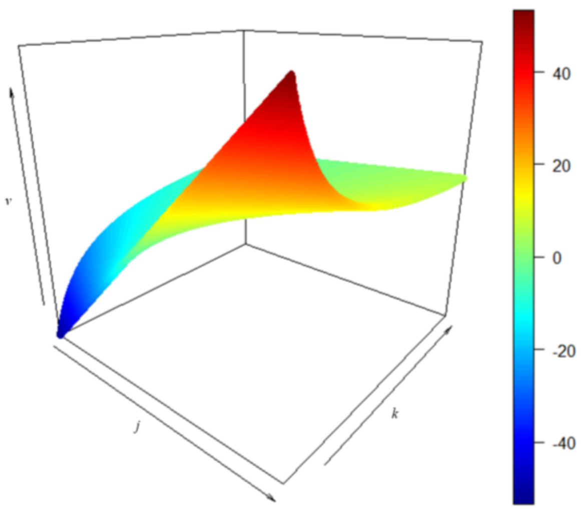



We will first explore the issue of convergence for different starting values under the same parameters. By setting the same parameters (simply set $j = k = 1$), as sequences starting from different x will eventually converge to the same value, we can plot the convergence process by repeating this operation 1024 times, as illustrated in the figure above. After 120 iterations, the convergence properties are already very intuitive, as all curves have converged to a point – in this example, all values will eventually converge to *(e/pi)/(1-(e/pi))*, which is around 6.4214796.




In fact, when the initial value $x$ and parameters $j$ and $k$ are different, the convergence shape of the sequence is not the same. We can set different parameters $j$ and $k$ to obtain convergence processes with different characteristics. For example, as shown in the above figure, when fixing the initial value of $x$ to 10000 and adjusting the values $k$ ranging from 0.2 to 10, we can observe that the rate of change exhibits different paths with different shapes. Thus, various convergence processes with varied properties can be simulated by different values of $j$ and $k$. Moreover, although the convergence process is discrete, we can obtain sequences with continuous properties using two methods. First, we can decrease the value of $k$, but be aware that this changes the convergence pattern. Second, we can use artificial neural networks to fit discrete sequences and obtain an approximate continuous model with continuous inputs and outputs.

…




In fact, the final convergence result of our iterative model is completely independent of the starting position $x$, as we have shown that as long as the iteration step n is large enough, any given finite real number x will eventually converge to a number with an expression of $(j*(e/pi)^k)/(1-(e/pi)^k)$. Therefore, we further consider only the final result of convergence after iteration, treating constant parameters $j$ and $k$ as two-dimensional vectors. By varying the parameter values in these two dimensions, we can plot the function surface of the convergence values as the parameter values are varied, as illustrated above. Please note that in the provided 3D graph, axis $j$ represents values of $j$, axis $k$ represents values of $k$, and axis $v$ shows the converged value of every combination of parameters $j$ and $k$, which is $(j*(e/pi)^k)/(1-(e/pi)^k)$. As shown in the figure, we find that the final convergence value $v$ is positively correlated with the parameter $j$ and negatively correlated with $k$. Moreover, the positive correlation between $v$ and $j$ mentioned above is linear, while the negative correlation between $v$ and $k$ is otherwise nonlinear, which is in full agreement with the properties of the convergence value expression we derived. In fact, by varying the values of parameters $j$ and $k$, we can obtain a convergent sequence that converges to any non-zero position, as the value of $j$ can either be positive, or negative.

Due to the unpublished status of this study, the effects of other properties of the model, such as $j$ and $k$, on the rate and location of convergence, or the properties of the convergence process will not be detailed illustrated. In terms of other important properties, we find that the convergence rate can become slower when applied with a smaller $k$, while it is faster for a larger $k$; variations of the joint parameters $j$ and $k$ can be made such that the converged value takes any finite real value; and for small values of $k$, the changing rate in the initial iteration step is approximately linear.

…

Due to the monotonically decreasing absolute value of the discrete rate of change in this model, its sequence values themselves have a similar "dissipative" effect. Therefore, a typical use of this model is to use convergence sequence superposition to simulate the new expected convergence trend of the economic market toward each shock after continuous or discrete shock reactions (of course, other random terms with arbitrary properties can be added outside the trend term to bring a certain degree of randomness). We assume that $P$ is an economic variable affected by discrete shock responses, and its sequence value in each period can be expressed in a general formula as follows:

$$P_n = \bar{C}$$

$$P_{n+a_1+a_2+\cdots+a_m} = f_{a_1}(X_{1_{a_1}}) + f_{a_2}(X_{2_{a_2}}) + \cdots + f_{a_m}(X_{m_{a_m}}) + \varepsilon_{n+a_1+a_2+\cdots+a_m}, which\ satisfies$$

$$X_{i_{a_{i+1}}} = (X_{i_{a_i}} + j_i) \times \left(\frac{e}{\pi}\right)^{k_i} + \epsilon_{i_{a_i}}, \lim_{iter:n\to\infty} \epsilon_{i_{a_{i+n}}} = 0, i = 1,2,\cdots,m$$

$$X_{i_m} = 0, \forall m < 0$$

Among them, $P_{n+a_1+a_2+\cdots+a_m}$ represents the value of the economic variable $P$ in the discrete shock response period of $n + a_1 + a_2 + \cdots + a_m$, $X_{i_{a_i}}$ represents the iterative sequence value with the parameter $j_i$、$k_i$ in the period $a_i$ using the above iterative model, $f_{a_1}$ represents a post-processing function that performs specific processing on iterative sequence values, and finally, $\varepsilon_{n+a_1+a_2+\cdots+a_m}$ is a random term. For the iterative value of $X_{i_{a_i}}$, Due to the fact that our model always has the property of eventually converging under a given parameter, we can set the next term of $X_{i_{a_i}}$ not only to be obtained by iteration, utbutlso to be affected by another random term $\epsilon_{i_{a_i}}$. As long as we ensure that the random term eventually converges to 0, the convergence properties of $X_{i_{a_{i+n}}}$ will be ensured without any change. Thus, in summary, in this general model, the economic variable $P$ has an aggregated tendency to be influenced by trend factors that start from different time periods, have different effects, and eventually converge to different values, while also having a certain degree of randomness.

Let us look at a simple example. One morning in early April 2023, AAPL announced at a press conference that a novel generation of smartphones had achieved a ground-breaking breakthrough in imaging technology that was expected to lead to a 15 percent surge in sales. The positive news was expected to drive the stock price to an expected higher value amid the volatility. In the immediate aftermath of the announcement, market prices began to adjust. But 45 minutes after the release of the press conference news (when the changes in AAPL's stock price had not fully absorbed the positive news released by the press conference - that is, the stock price was still in the process of adjustment), the Federal Reserve released a new interest rate hike of 75 BP, without warning, which is then expected to bring about a short-term bearish trend in the entire securities market (due to a decrease in the expected excess returns of risky assets), and this would indeed, cause the stock price move towards an expected lower value amidst fluctuations.

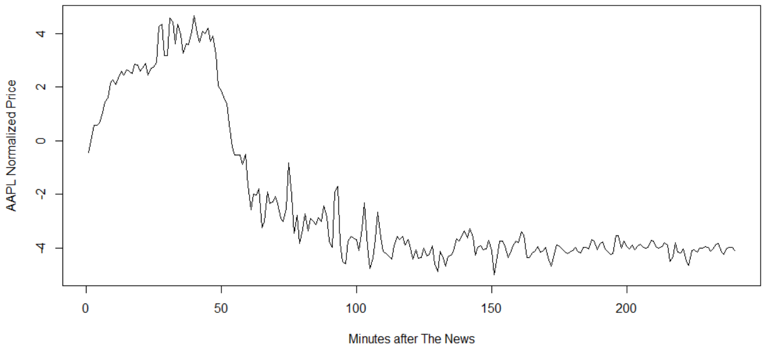
By simulating the above two shocks, we assume that time is discrete (such as in minutes), and assuming that there is no other trendy stock price influencing factors during this process, we can establish an iterative model with the following shock response form:

$$P_n = \bar{C}$$

$$P_{n+a_1+a_2} = f_{a_1}(X_{1_{a_1}}) + f_{a_2}(X_{2_{a_2}}) + \varepsilon_{n+a_1+a_2}$$

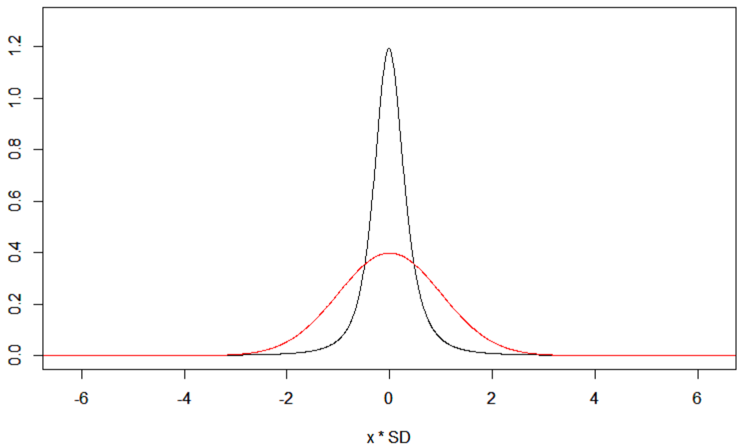
$$X_{i_{a_{i+1}}} = (X_{i_{a_i}} + j_i) \times \left(\frac{e}{\pi}\right)^{k_i} + \epsilon_{i_{a_i}}, \lim_{iter:n\to\infty} \epsilon_{i_{a_{i+n}}} = 0, i = 1,2$$

$$X_{i_m} = 0, \forall m < 0$$




In fact, given the specific forms of the above parameters and functions, we can draw a random sequence of changes of 240 minutes in normalized stock prices (assuming that the starting price was 0) after being impacted, as shown in the following figure. We find that after two shocks happening at 0 and 45 minutes, the normalized stock price of AAPL eventually stabilized at -4 without any other effect. Please note that this simulation is only for demonstration, since it has a very simple form of the random term, and in real application, the random term can satisfy an ARMA or other forms of time-varying stationary sequences.

…




We have also explored a particular use of the above model when fitted as a continuous function. Due to the arbitrariness of the choice of parameters j and k in the proposed model, it is possible to obtain a bell-shaped distribution of the rate of change of the iteratively generated discrete sequence after obtaining the rate of change of the sequence at a certain parameter and initial position. We used the random number sequence generated by the above distribution to draw a frequency distribution histogram (i.e. PDF of the distribution) and observe its statistical characteristics. We find that it features a mean of 0, a standard deviation of 0.5, unbiasedness, and a kurtosis of 6.88. That is, the largest feature of this symmetric distribution is its super-peak character, with about 78 percent of the data distributed within a plus or minus two standard deviations. As illustrated in the plot above, while the red line shows the PDF of the standard normal distribution, the black line shows the new, super-peak distribution.

## 3. [Unpublished Trading Strategy Research] An Active Bond Trading Strategy Based on An LSTM Neural Network Momentum Inverse Model

**Description**--------------------------------------------------------------------------------------------------------------------

Status: unpublished professional results obtained when doing an internship

Proposed Time: June, 2022

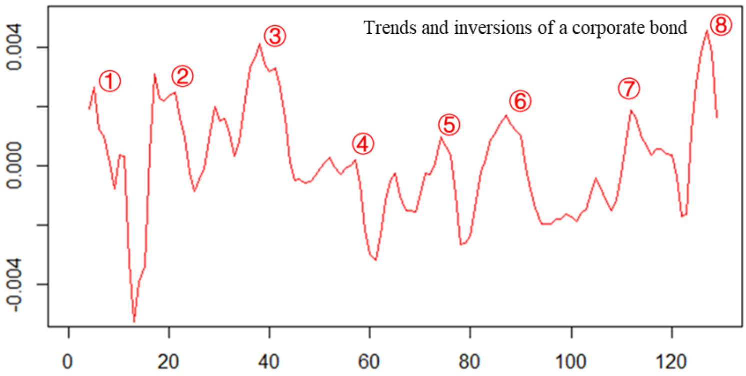
Interest rates are theoretically a killer factor in pricing a bond and an important element in influencing the market price of a particular bond. In our study, starting with an aim to capture the momentum effect of interest rates and bond prices, we constructed a neural network interest rate prediction model based on LSTM and a trend prediction model by using a micro-predictive approach and then combined them into a TECH active trading strategy. For the data applied, 140 coupon-bearing treasury bonds, local debt, corporate debt, and corporate debt issued and circulated in 2020-2022 were used for back-testing experiments, and 40 bonds were then plotted with a CIAV trading record indicating the stability in its trading process. It found that compared to control groups namely YTM and HOLD strategies, our TECH trading strategy is able to achieve superior excess returns and higher Sharpe ratios by forecasting and warning to capture upward trends and resist downward ones, having effectively grasped the impact factors of interest rates on bonds and provided meaningful warnings in time when trends are predicted to reverse. Regression tests demonstrate the remarkable ability of TECH's active trading strategy to capture excess returns by capturing trends in volatility.

…

**Basic Principles of Interest Rate Forecasting**------------------------------------------------------------------------------

…

Interest rates are influenced by macroeconomic factors such as GDP, savings rate, and monetary policy, and therefore the common models used by a large number of institutions for interest rate forecasting are normally based on those macroeconomic factors. Although predicting interest rates based on macro factors is relatively more available due to the frequency and difficulty of quantification of macroeconomic data, nevertheless, it often fails to achieve a high frequency within that framework (such as weekly and daily interest rate forecasting).

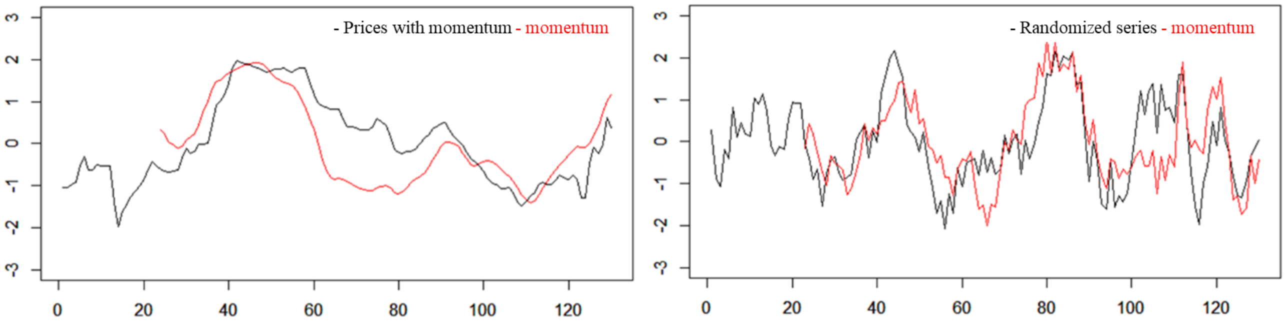



In reality, bond prices are generally subject to higher frequency changes in interest rates, particularly for various types of corporate bonds and company notes, where the trend of price increases and decreases is usually sustained only at daily levels, for example, an increasing trend that lasts for at most 2 weeks. As a result, low-frequency interest rate forecasts based on macroeconomic factors can hardly satisfy the requirements of bond trading, which typically takes place daily. The figure illustrated above shows the trend change in the price and yield of a particular Chinese corporate bond. The number of trend changes was estimated to be approximately 7.69 by using the run test of a 100-day sample processed by an MA process with 4 lag phases, which also means an average persistence time is about 13 days for each upward or downward trend.

…

The momentum effect was proposed by Jegadeesh and Titman (1993), whose research results indicated that stock returns tend to continue their original direction of motion, and so far, numerous empirical studies have explicitly shown that many financial variables have trends implied by momentum, such as bond prices, futures prices, exchange rates, interest rates, etc. Let's start by exploring the momentum of bond prices.

A simple way to test the strength of momentum effects is to do the run test based on the momentum factor. The momentum factor can be calculated based on the ROC index, which can calculate the momentum level of a time series at a certain time point under a frequency - for example, a momentum index of 0.3 indicates that the time series has an upward trend at that time point. The criterion by using the momentum factor is based on the fact that time series data with significant momentum will exhibit a continuous positive or negative momentum for a certain period when the decomposition frequency increases, at which time the momentum indicator will continue to increase or decrease. In contrast, time series data without significant momentum will not show significant momentum as the decomposition frequency increases, so the rate of change of the momentum indicator will continue to fluctuate through the zero line.

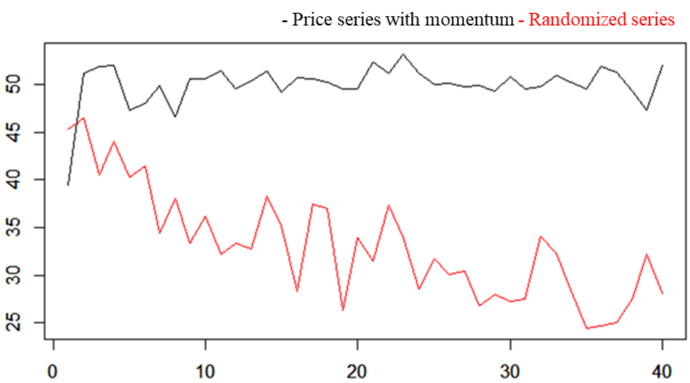



For example, the figure above implies the real price fluctuations of a certain bond and the momentum indicator, as well as a random walk time series for comparison. It can be observed that the bond's momentum indicator moves with more trend patterns, while the growth rate continues to fluctuate and is more randomized.

…

As illustrated above, results of the run test (vertical axis, %) of the momentum indicator of the bond price time series show a significant downward trend with the increase of decomposition frequency (horizontal axis), and quantitatively, this significant trend can be determined by using an OLS linear regression. For example, the following model can be used to fit the values of the run test results mentioned above.

$$reg: run(Mtm) \sim \alpha + \beta_1 f + \beta_2 f^2 + \varepsilon$$




The reason why it takes the form of a mixture of linear and quadratic patterns is that we observe a significant downward convexity in the downward trend, as shown in the figure above. Using the bond data we demonstrated previously for estimation, the regression coefficients $\beta_1$ and $\beta_2$ are -0.857 and -0.011 respectively, and the t-values are -4.725 and 2.653 respectively, both of which are significant at the 95% confidence level, hence, it can be judged that the bond price time series has momentum. By generalizing this approach, we can use the regression model described above for a larger testing area. For any bond to be tested, the first step includes calculating its momentum sequence, then using the above regression equation to estimate it, and finally, a decisive factor lies directly in the significance of the coefficients $\beta_1$ and $\beta_2$ (we have to say that normally, the significance of the coefficients $\beta_1$ can be more important since it directly indicates the 'downward' trend that defines the significance of momentum rather than simply implying a form with convexity).

We randomly selected 130 bond price time series from the sample bank to perform the aforementioned run tests on momentum factors and performed the aforementioned regression on the results. The t-value levels for $1\ for$ all bonds are shown below. It is observable that almost all bond prices have significant momentum under our regression significance framework.

…

**LSTM Model for Interest Rate Forecasting**--------------------------------------------------------------------------------

…

We used SHIBOR's 1-month period of 2400 daily data for differential and momentum calculations to complete a training dataset of 2400 samples, with an additional 600 samples serving as the test set. After training at a learning rate of 0.0001, the final LSTM neural network model was obtained after 20000 iterations. In contrast to the prediction accuracy of linear regression models constrained by the form of the model, as for our LSTM model, its memory characteristics and nonlinear fitting ability appear to be an outstanding advantage to finding patterns in complex data, but scientifically, we also need to evaluate the prediction performance of LSTM models. As for the evaluation factors, after rounds of experiments performed, we have selected common MSE, RMSE, MAE, and CORREL indicators for evaluation, and for MSE, RMSE, and MAE, which are common indicators for measuring the error of two sets of data, the smaller their values, the better the prediction accuracy, while CORREL is a linear correlation coefficient, which means that the larger its value is, the better the prediction accuracy will accordingly be. The evaluation approach above will be performed on the test set mentioned above, with 600 samples, and all results generated from both methods will be compared between the LSTM model and the linear regression model.

| | MSE | RMSE | MAE | CORREL |
|---|---|---|---|---|
| lm | 0.0356 | 0.189 | 0.104 | 0.2823 |
| LSTM | 0.0003 | 0.017 | 0.016 | 0.9998 |

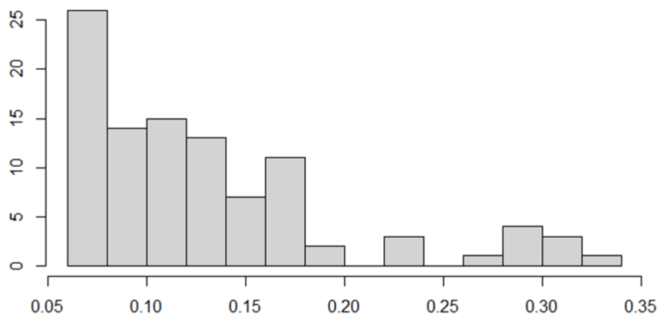
Take some insights into the statistical test results. In terms of error indicators like MSE and RMSE, the performance of the LSTM model is significantly better than that of the LM model. Among them, the RMSE error of the LSTM model on interest rates is 0.017%, while that of the LM model is 0.189%. In addition to the error factors, considering CORREL that indicates relevance, the correlation coefficient of the LSTM model is 0.9998, while the LM model correspondingly shows only 0.2823 of its linear correlation. It is remarkable that these significant differences truly reflect the ability of neural networks with long short-term memory properties to fit time series with temporal features, as well as their memory's ability to characterize trends.

…

We have proposed a new framework to evaluate the performance of the LSTM model in this specific interest rate prediction task. Since we already have two price series, which are the real one and the predicted one based on the lag data, we can directly use the accuracy of trend prediction (TRP) to measure the LSTM model's ability to depict trends by describing the difference between the two series. If the length of the sequence is n, TRP can be expressed as the following formula:

$$TRP = \frac{1}{n}\sum_{t=1}^{n} slope(prediction) \times slope(price) < 0\ ?\ 1 : 0$$

The TRP is calculated as a scalar quantity that measures the ratio of all predicted values to the actual values that change in the wrong direction. Thus, it also indicates that the lower the TRP value of a predicted sequence, the more it implies that the predicted sequence is closer to the actual sequence in terms of the direction of change – which, of course, is in accordance with our needs for interest rate prediction – to predict the direction.




We used the structure of LSTM to model and predict different types of interest rates to further observe the average TRP performance other than the performance from a single sample set. Across dozens of data types, the average TRP of the LSTM model is approximately 13.18%, and the median is 11.09%, which all together, means that in most cases, the trend of interest rates can be correctly predicted by the LSTM model. The figure above shows the distribution of all TRP values computed from a variety of interest rate data, and observably, most cases tested have illustrated a relatively low level.

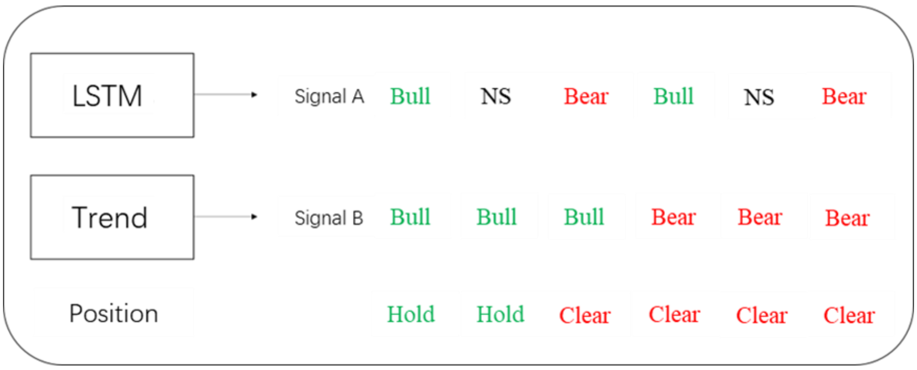
…

Note that the part of building up a trend prediction model based on the momentum effect has been truncated.

…

Based on the previous research, we have established an interest rate prediction model based on an LSTM neural network and a trend prediction model based on the momentum effect, which can all predict the future price trends of bonds from different perspectives. However, due to the different characteristics of the two separate models, the LSTM model can fundamentally predict the changes in bond value with changes in discounted interest rates, but the trend prediction model can further confirm the direction of bond price changes based on the momentum factor generated by bond prices. Needless to say, the latter can be efficiently used mainly to avoid the decoupling of interest rate changes and price changes when incorrect pricing occurs, so it is necessary to concurrently perform both models and use them complementarily.

In terms of the specific complementary use, we have defined a logical operation on the prediction results of the two separate signals to jointly make the final prediction decision. Please note that Chinese bonds cannot be shorted, so when the short-selling signal appears, we choose to sell and not to hold it anymore, and accordingly, this period without position will not be included in the annualized rate of return (because we can still choose to hold other bonds or financial products to create revenue). The method of logical operation is as follows:




Our strategy TECH relies on two kinds of signals obtained from the Trend model (respectively, "Bull", "Bear"), and three kinds of signals from the LSTM model (respectively, "Bull", "Bear", "NS"). The processing process is similar to a logical operation of "AND", that is, the final decision will only be held when the signals of both models indicate "Bull", or the trend model indicates "Bull" and the LSTM model has a "Not Specified" signal. The remaining cases simply imply a short position. We can assert that this is a more conservative strategy since we will only take a long position when and only when a very strong signal has been confirmed by two separate, robust models.

…

We have also set up two other strategies as a control group to test the effectiveness of our proposed TECH strategy. One inside the control group is a "YTM Oriented Trading Strategy" (noted as YTM), which will only take up a long position when the market price is regarded as undervalued and sells the bond when overvalued. Additionally, another strategy is the "Passive Continuous Hold Strategy" (noted as HOLD), which will simply hold the strategy from start to end, simulating the returns under passive management.

It should be reiterated that our back-testing experiments will use, and only use clean prices since they do not include accrued interest that might significantly differ from bond to bond due to the different coupon rates they have. Therefore, the investment income from clean price transactions can be indeed deemed as the "excess revenue" brought by a kind of "interest rate speculation". Furthermore, the durations of all bonds we included lie between 1 and 10 years to avoid legacy bonds with outlier interest rates, and statistically, obey a normal distribution at the 99% confidence level.

…

**Effectiveness Analysis of The Strategy** ------------------------------------------------------------------------------------

| | | mean | mid | sd | var | skew | kurt | t val |
|---|---|---|---|---|---|---|---|---|
| TECH | Achieved Return | 0.0203 | 0.0195 | 0.0137 | 0.000188 | 1.3010 | 2.7394 | 1.4811 |
| | Annual Return | 0.0595 | 0.0422 | 0.0491 | 0.002415 | 1.4071 | 1.1480 | 1.2110 |
| | Annual SD | 0.0152 | 0.0119 | 0.0102 | 0.000104 | 0.5989 | -0.9799 | 1.4881 |
| | Sharpe Ratio | 4.4722 | 4.3056 | 2.5608 | 6.5578 | 1.2701 | 2.1387 | 1.7464 |
| | Chance Ratio | 0.1926 | 0.1722 | 0.1010 | 0.010210 | 1.4440 | 1.8873 | 1.9062 |
| | Value of Risk 95% | -0.0010 | -0.0008 | 0.0008 | 0.000001 | -0.9344 | -0.1263 | -1.3045 |
| | Ex-Achieved Return | 0.0170 | 0.0168 | 0.0142 | 0.000201 | 1.2588 | 3.6326 | 1.2024 |
| | Ex-Annual Return | 0.0519 | 0.0328 | 0.0432 | 0.001870 | 1.1843 | 0.6851 | 1.2004 |
| YTM | Achieved Return | 0.0012 | 0.0003 | 0.0150 | 0.000225 | 0.1557 | 0.7353 | 0.0774 |
| | Annual Return | 0.0042 | 0.0004 | 0.0599 | 0.003583 | 0.2437 | 1.6592 | 0.0697 |
| | Annual SD | 0.0149 | 0.0124 | 0.0097 | 0.000095 | 0.6616 | -0.5043 | 1.5322 |
| | Sharpe Ratio | 0.4300 | 0.0571 | 3.3303 | 11.0907 | 0.6070 | 0.5030 | 0.1291 |
| | Chance Ratio | 0.0204 | 0.0027 | 0.1586 | 0.025146 | 0.3550 | 0.7678 | 0.1288 |
| | Value of Risk 95% | -0.0013 | -0.0010 | 0.0010 | 0.000001 | -0.9124 | -0.0006 | -1.3410 |
| | Ex-Achieved Return | -0.0021 | 0.0000 | 0.0148 | 0.000219 | -1.1276 | 1.4895 | -0.1446 |
| | Ex-Annual Return | -0.0034 | 0.0012 | 0.0522 | 0.002722 | -0.6486 | 0.6625 | -0.0656 |
| HOLD | Achieved Return | 0.0033 | 0.0029 | 0.0125 | 0.000157 | 0.4936 | -0.1688 | 0.2636 |
| | Annual Return | 0.0076 | 0.0038 | 0.0225 | 0.000507 | 1.1058 | 1.9694 | 0.3374 |
| | Annual SD | 0.0218 | 0.0189 | 0.0120 | 0.000145 | 0.1392 | -1.3069 | 1.8090 |
| | Sharpe Ratio | 0.4216 | 0.1467 | 1.3623 | 1.8558 | 1.2053 | 2.4463 | 0.3095 |
| | Chance Ratio | 0.0354 | 0.0220 | 0.1713 | 0.029331 | 0.0241 | 1.0801 | 0.2064 |
| | Value of Risk 95% | -0.0020 | -0.0018 | 0.0012 | 0.000002 | -0.3711 | -1.0216 | -1.6462 |
| | Ex-Achieved Return | - | - | - | - | - | - | - |
| | Ex-Annual Return | - | - | - | - | - | - | - |

For the back-testing experiment we conducted, we used the NKFQ automatic trader to perform a pure automatic planning transaction based on the logical operation values of two signals generated by our LSTM and trend models. Please note that we assume that there is no market friction, that traders are price takers, and that there are no taxes or fees that might have devalued our revenues (although it is slightly idealized). All trading operations based on historical data will only be used as instructions for the next period's trading. All transaction data have already been cleaned, time aligned, and converted to net prices (excluding coupon returns) before performing the back-testing experiment, so the returns only include the net returns of "interest rate speculation". The automated trader verifies all the given data, but in the first 20 samples it only verifies and does not trade, while starting from the 20th sample, transactions will officially begin.

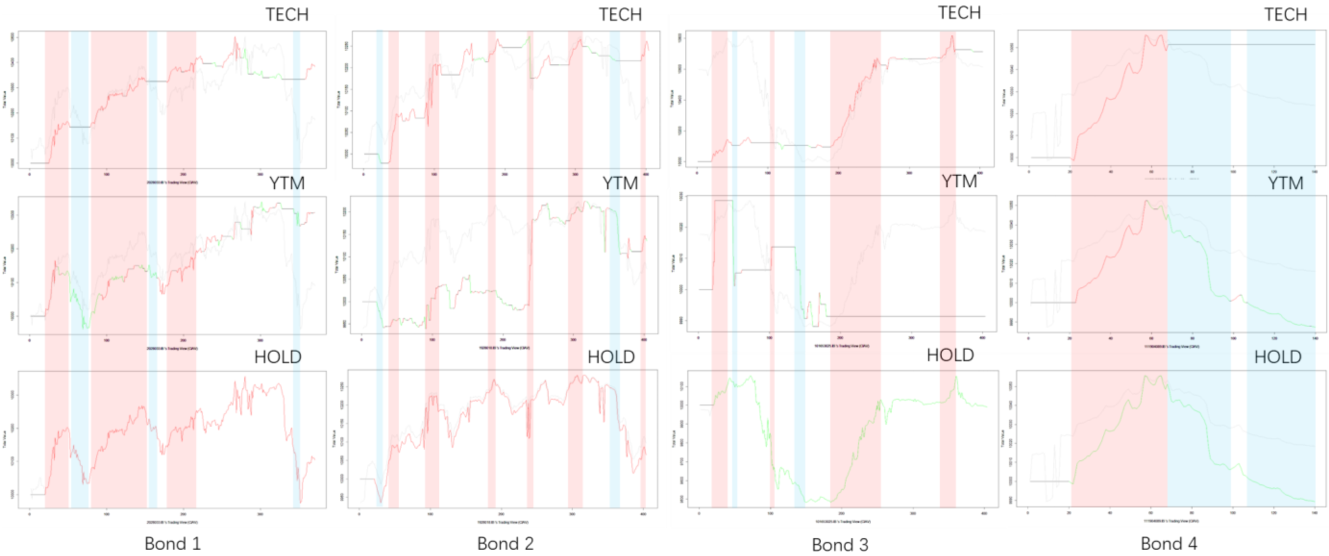
We conducted back-testing trading experiments on the daily clean price series of 670 Chinese coupon bonds using the TECH, YTM, and Γ strategies, respectively and provided a full statistical table of the trading results as shown above.

…

First, look at the annualized returns of different strategies. Due to the different maturities of the bonds, we need to convert them to annualized returns for easier comparison of different strategies, hence, both directly obtained returns and annualized returns are shown in the table. Our proposed active trading strategy TECH achieved an average annualized return of 5.95% and a median return of 4.22%. In contrast, the YTM strategy and HOLD strategy achieved average annualized returns of 0.42 percent and 0.76 percent, respectively. In terms of annualized returns, the active strategy TECH achieved significant excessive returns compared to other strategies since the difference in its data has a significant meaning at a 99% confidence level. Note that the annualized return data has a right-skewed character, as the bond net price also shows an upward trend as the maturity date approaches. Therefore, the likelihood of any strategy achieving a significant negative return in bond clean price trading is very low, resulting in a right-skewed yield. In summary, our active trading strategy TECH can effectively predict interest rates and trends to avoid holding bonds when prices drop significantly, thus further increasing the yield skewness property to the right.

…

The above analysis indicates that our proposed trading strategy TECH has achieved a significant advantage over the YTM and HOLD strategies in the control group in terms of annualized returns and corresponding risks. Now, it is time to take deep insights into the reason why those differences existed. We can use CIAV figures for analysis (CIAV is a visual plot of the trading behavior of automatic traders under the NKFQ analysis framework in R) - to analyze what the YTM strategy was doing when the TECH strategy gained profits, and what the TECH strategy was doing when YTM strategy generated losses, to further explain the underlying story of the gained excessive returns.




(Note that this plot is illustrated with Chinese color preferences, with RED for up and GREEN for down.)

As shown in the combined CIAV plots above, we selected 4 bonds with different price characteristics from 670 bond samples, among which bonds 1 and 2 have a fluctuated, upward trend, bonds 3 and 4 have a V-shaped fluctuation trend, bonds 2 and 3 have short-term profit opportunities, and bonds 1 and 4 have long-term upward profit opportunities.

Looking at the combined CIAV plots, we can see that the TECH strategy mainly relies on trend profitability, while the YTM strategy mainly relies on mispricing profitability. The main reason why the overall profitability of the TECH strategy can outperform that of the YTM strategy is that the profitability of bond price movements is much stronger (for the existence of momentum effects) than that of mispricing, which strongly requires the market to fix it in a short period of time. However, if the mispricing cannot be fixed in the near future, applying the YTM strategy will incur significant losses and might be indeed forced to clear its position in reality.

The case where the TECH strategy is able to capture the future upward profit from the upward trend promptly can be obtained from the CIAV images, as shown in the first red color block for Bond 2 and the third red color block for Bond 3 in the above figure. In addition, the TECH strategy is able to sell and short positions promptly after predicting future downward trends, reducing losses caused by significant price drops, as shown in the first blue color block of Bond 1, the second blue color block of Bond 3, and the first and second blue color blocks of Bond 4 in the figure above. While the TECH strategy is short, the YTM strategy may still hold positions due to mispricing that has not yet been fixed, resulting in reduced returns or even losses due to price declines compared to the robust, stable returns that our TECH strategy has achieved.

…

## 4. [A Published Paper] Theoretical Analysis and Empirical Research on the Supply and Demand Relationship of US Stocks and the Long-term Growth Trend of Stock Index

**Description**--------------------------------------------------------------------------------------------------------------------------

Status: published paper

Proposed Time: April, 2022

Special Note: This paper was jointly completed by the applicant himself and the Professor. Zhijun, Wang.

Since the 1980s, the U.S. stock market has shown a long-term, sustained uptrend. There are many reasons that can explain this phenomenon, but among them all, the fundamental roles of the supply and demand relationship of listed stocks fundamentally contributed to this phenomenon. Statistical data show that in the past forty years, the demand for listed stocks in the United States has continued to grow, while the supply of listed stocks has continued to decrease, and under the combined effect of demand and supply side forces, the stock prices of listed companies in the United States have continued to rise. This paper uses the Federal Reserve's data on the issuance, purchase, and related economic data of US-listed stocks from 1985 to 2020 to empirically analyze the relationship between changes in supply and demand of US-listed stocks and long-term stock index rise, with the aim of contributing to the understanding of the development of the US stock market.

The main contributions of this paper are listed as follows. First, this paper proposed a set of driving forces that affect the sustained growth of the US stock index from the perspective of supply and demand in the stock market and conducted a theoretical analysis of the three major indicators using supply and demand models to determine the transmission path that leads to the sustained growth of the U.S. stock index under supply and demand factors. Second, based on the hypothesis of significance for the three major factors mentioned above, this paper constructed a macroeconomic model of supply and demand in the stock market, and empirically tests the three major factors that affect the sustained growth of the U.S. stock index, confirming their significant role in stock index growth with macroeconomic data of the U.S.

…

**Characteristics of Supply and Demand Relationship in Contemporary Stock Market of the U.S.** ---------------

**Chr I. The high securitization rate of household financial assets in the United States.**

**Table 1** Financial Asset Securitization Ratio of US Household Investors (100 thousand)

| | 1985 | 1990 | 1995 | 2000 | 2005 | 2010 | 2015 | 2020 |
|---|---|---|---|---|---|---|---|---|
| Financial Assets | 10864.2 | 15950.5 | 22766.6 | 34313.8 | 46166.5 | 53810.9 | 70782.0 | 104767.3 |
| Securitized Assets | 5718.4 | 9531.6 | 15826.6 | 25017.3 | 31237.2 | 38477.6 | 49499.4 | 75805.2 |
| FASR (%) | 52.64 | 59.76 | 69.52 | 72.91 | 67.66 | 71.51 | 69.93 | 72.36 |

*Source: Federal Reserve, Financial Accounts of the United States. L100, etc.*

U.S. household investors, while investing directly in securities, typically invest more of their investment and wealth management funds into institutional investors for investment, retirement, and risk-prevention purposes, thereby indirectly owning significant amounts of securities, particularly publicly traded stocks. Quantitatively speaking, the preference of American household investors for holding securities can be measured using the Financial Asset Securitization Ratio (FASR), which can be calculated by the ratio of securitized financial assets to all financial assets those households own. As shown in Table 1, the FASR of US households and individual investors has been continuously staying above 50% for the whole period of time starting from 1985 to 2020. Only in 1985 and 1990 did the FASR fall below 60 percent, but, in particular, since 1995 it has returned to its original level and stayed consistently above 60 percent, although depicting a slightly fluctuating trend.

U.S. household investors have a generally higher securitization ratio, which means a significant preference for financial assets compared to holding other kinds of assets. When national income and marginal investment propensity are fixed, U.S. households and individual investors will use more of their funds to invest and hold stocks, directly or indirectly. Reflecting on the stock market, stocks directly held by American household investors have directly become an effective demand for listed stocks in the United States, while the indirectly held portion constitutes the source of funds for financial institutions in the form of financial instruments, altogether laying the foundation for their stock demand.

…

**Chr II. The increasing degree of institutionalization of stock investors in the United States.**

**Table 2** Shares held by major shareholders of US-listed companies (%)

| | 1946 | 1960 | 1970 | 1980 | 1990 | 2000 | 2010 | 2020 |
|---|---|---|---|---|---|---|---|---|
| Market Value (Billion $) | 109.7 | 420.3 | 831.2 | 1494.9 | 3531.3 | 17627.0 | 22961.6 | 64502.6 |
| Household | 92.3 | 85.6 | 78.2 | 67.6 | 55.5 | 45.9 | 35.9 | 38.3 |
| International Institution | 2.5 | 2.2 | 3.2 | 5.0 | 6.9 | 9.3 | 13.4 | 16.3 |
| Private Pension Fund | 0.3 | 3.9 | 8.1 | 15.5 | 17.2 | 10.9 | 8.8 | 5.4 |
| State & Local Pension Fund | -- | 0.1 | 1.2 | 3.0 | 8.1 | 7.4 | 7.8 | 4.6 |
| Mutual Fund | 0.9 | 3.5 | 4.8 | 2.8 | 6.6 | 18.3 | 20.7 | 20.8 |
| Closed-end Fund | -- | 1.2 | 0.6 | 0.3 | 0.5 | 0.2 | 0.4 | 0.2 |
| Exchange Traded Fund | -- | 0 | 0 | 0 | 0 | 0.4 | 3.7 | 6.6 |

*Source: Federal Reserve, Financial Accounts of the United States. F213, etc.*

In 1946, household investors in the United States held 92 percent of the total market capitalization of publicly traded U.S. companies, according to statistics from the U.S. Financial Accounts, meaning that the market at that time could be considered fully retail. With the post-war US market getting mature and advanced, however, this proportion had dropped to 68% by 1980 and had then sharply declined to 38% by 2020 (as shown in Table 2). In the meantime, by contrast to the decline of the proportion that directly holds stock shares, the proportion of shares held by financial institutions and international investors has increased from almost zero to nearly half, more than the level of individual holdings, representing a continuous enhancement in the process of institutionalization. Please note that institutionalization of investors in the US stock market does not only refer to domestic financial institutions in the U.S. but also includes foreign investors since foreign investors can only invest those US-listed stocks through international financial institutions as specified by the regulations.

…

Turning our focus on capital flows, the household sector generally had a negative net annual direct stock investment over the past four decades, meaning they have sold more of their stock than bought, but instead, we found that a significant amount of capital from the household sector has been flowing into pension funds, mutual funds, life insurance companies, and any other financial institution, meaning that indirectly holding equities through these institutions have become a permeated choice. According to Federal Reserve statistics, the indirect ownership of stocks by domestic investors has strengthened American institutions and contributed to the institutionalization of the US stock market.

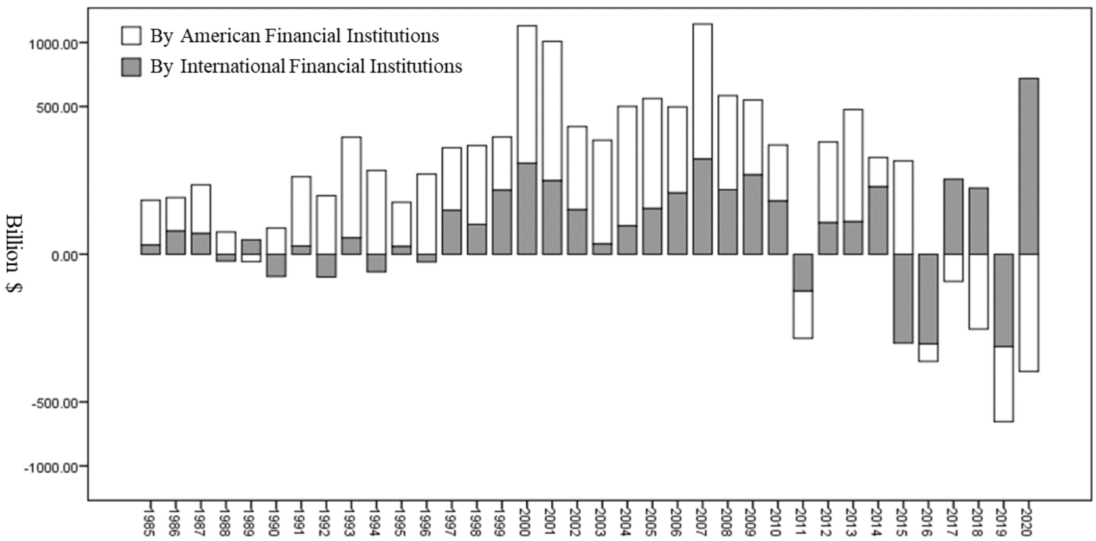
…

**Chr III. The stable growth in total stock demand of listed stocks in the United States.**

The aggregate demand sector in the U.S. stock market for publicly traded stocks can be divided into demand from households and financial institutions. With the trend of an increasing level of securitization and institutionalization, the direct and indirect stock demands from the household sector have been driving continuous growth in the overall stock demand, driving up the aggregate demand at the same time.

First, investment demand from the household sector in the United States has illustrated long-term stability throughout history, and particularly, the preference for securitization of financial assets of households has laid the overall foundation for the giant, everlasting robust demand. In particular, given the same amount of disposable income, US household investors allocated a significantly larger proportion of their investments to securities assets than in other countries, which means a greater portion of US households and individual investors preferred to hold securitized assets, particularly equities, rather than other assets such as bonds and bank savings, regardless of macroeconomic fluctuations during this long period of time. For instance, even in the year with the highest institutionalized rate in 2010, the U.S. household sector still directly held approximately 35% of the market value of listed stocks, making up an electable portion to be analyzed.

**Figure 1** Annual net purchases of listed stocks by US and international financial institutions




*Source: Federal Reserve, Financial Accounts of the United States. F213, etc.*

Second, demand from institutional investors has been steadily increasing as well. Statistical data from the Federal Reserve indicates that from 1985 to 2020, financial institutions maintained positive net purchases of stocks in almost every single year (referring to Figure 1), while only in a minority of years, data of net purchases from institutions show a negative number. As a result of their constant purchases, institutions have been gradually accumulating more and more equity assets over the years.

In addition to the large volume of demand, institutional investors in the US have a more stable investment profile. According to data provided in Financial Accounts of the United States, in an upward market cycle, when economic growth finally led to an increase in national income, which provided basic support for households and individual investors to purchase financial assets, household investors, through the purchase of financial instruments such as funds by financial institutions, continuously injected investment funds into financial institutions by indirectly purchasing financial instruments, which were then converted to direct demand of financial institutions. By contrast, in a downward cycle of the market, due to the mandatory investment policies, professional management abilities, and long-term prospects of those financial institutions, their purchases of stocks were observably less volatile to panic selling, and their demand shows more stability regardless of the market fluctuations.

…

**Chr IV. The continuous decline in the supply of listed stocks in the United States.**

Total supply in the US stock market can be divided into incremental factors and so-called existing factors, including all of those public and tradable stock assets. Among the two categories, Incremental supply mainly includes positive factors such as IPO and SEO that bring fresh blood to the market, while existing supply includes negative factors such as repurchases and delisting of listed stocks, which accordingly, control the market from being too extreme. The relative shrinkage of IPOs in the US stock market, repurchases of listed stocks by large companies, and structural delisting of small companies have led to a long-term phenomenon of negative net stock issuance, resulting in a continuous decrease in the total supply of stocks in the market.

On the one hand, the initial public offering (IPO) of stocks is the act of a listed company selling its stocks to the public for the first time, usually happening at the beginning of its listed period. Considering the necessity of financing such a great amount of money at this stage, IPOs always constituted the most important initial stock supply in the stock market and usually brought a giant number of fresh, tradable stocks into the market, therefore, holding a fundamental position in incremental supply for the stock market. In comparison with the booming period of IPOs after World War II, the number of those in the late 1990s otherwise illustrated a significant, irreversible downward trend. For example, in the most prosperous year of 1996, the number of IPOs in the US stock market reached a relative peak of 860, but since then, both the number and value of IPOs in the US stock market have significantly shrunk, leading to a decrease in incremental supply.

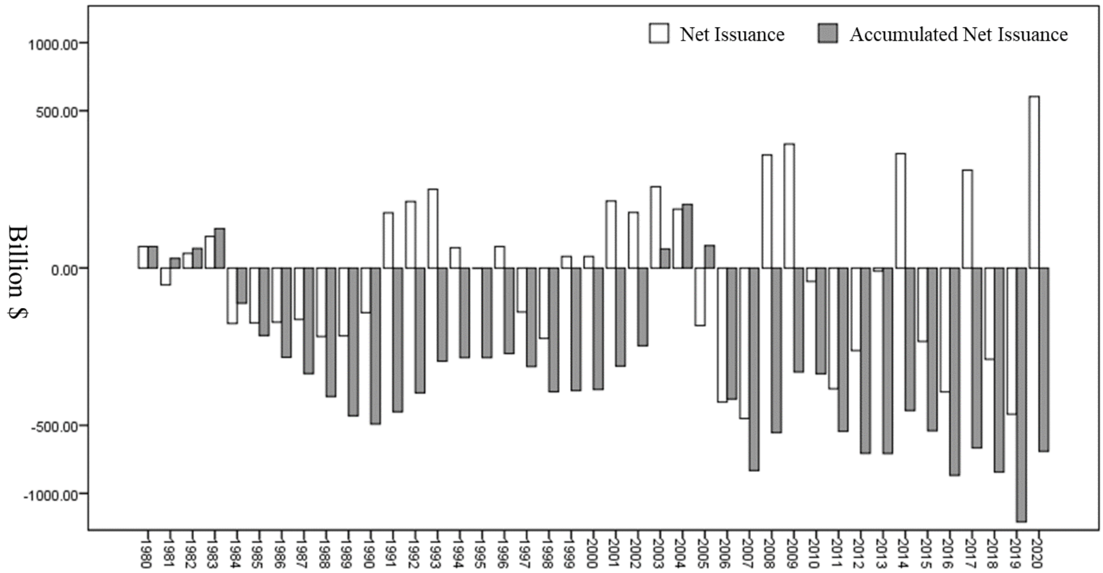
Coincidentally, against the background of the reduction of incremental stocks, the reduction of existing stocks has also become a market-wide phenomenon. Since the end of the 20th century, share repurchases have become an increasingly welcomed approach for U.S. public companies, especially those large, prominent ones with good cash flow, such as Apple and Coca-Cola, where they promote shareholders' interests in a way that can circumvent any cash dividends that will cause a shrinkage of stock prices. In recent years, the scale of repurchases in the U.S. stock market has further expanded. Quantitatively speaking, from 2000 to 2018, according to data from Financial Accounts of the United States, a total of 17,639 common stock repurchases occurred in the U.S. stock market, and the announced accumulated repurchase amount reached US$8.6 trillion at the end of 2018. It is worth noting that U.S. listed companies usually prefer to cancel the shares that they repurchased, to directly reduce the existing supply of the market to gain additional capital premium and push up the price. If the scale of repurchases and cancellations is relatively large, it may cause the net issuance of the entire market to become negative, leading to an observable reduction in the supply of shares.

…

As a result of the factors listed above, net equity issuance by listed companies has been a consistently negative number from 1985 to 2020 according to Federal Reserve statistics. In particular, U.S. corporate equity net issuance has been a negative number for 20 of the past 25 years, with a cumulative net issuance of $863.6 billion, hence, the negative net issuance can be considered a norm (as shown in Figure 2), reflecting the continuous decrease in the supply of listed stocks in the US stock market.

In summary, akin to general commodities, stocks, as a special kind of financial good, also follow a general law of value, where the price of a stock is defined by both its supply and demand. According to the efficient market theory, supplement relationships of stocks are the determining factors of their equilibrium prices, which, in particular, rise significantly when demand increases and supply decreases. In reality, since the era of the later 1980s, the steady increase in demand for listed stocks and the continuous decrease in supply has shaped and enhanced this relationship of supply exceeding demand, and as a result, the stock price index has predicted a continuous upward trend.

**Figure 2** Changes in net issuance and accumulated net issuance of US-listed companies




*Source: Federal Reserve, Financial Accounts of the United States. F213, etc.*

…

**Research Hypotheses and Empirical Testing --------------------------------------------------------------------------------**

…

In the theoretical analysis of the reason why the US stock index depicted a continuously upward trend, we proposed an explanation that the radical drive behind the increasing price lies in the relationship of supply exceeding demand, and in order to imperially test the validity of our explanation, particularly the effects that the demand and supply caused, the following two hypotheses are proposed:

**Assumption 1**: The sustained growth of total demand in the US stock market has a significant positive impact on the sustained growth of the US stock index.

**Assumption 2**: The continuous decrease in total supply in the US stock market has a significant positive impact on the sustained growth of the US stock index.

In terms of dependent variables, the stock price index, we obtained the US Dow Jones Industrial Average from the Wind database from 1985 to 2020 as the test stock index. Due to the characteristics of obtaining other macroeconomic data, the annual data is used as the unit of dependence variable in this paper, and the other explanatory variables are also using the annual data.

On the other hand, in terms of explanatory variables, according to the assumed conditions, the net increase in total demand for stock assets by households, US financial institutions, and international financial institutions is chosen as the explanatory variable to measure the sustained growth of total demand, and the cumulative net issuance of US stocks is chosen as the explanatory variable to measure the continuous decrease in total supply. In terms of specific composition, the net increase in total demand is represented by the net increase value of the total holding value of all investors' stocks, including household sectors, property insurance companies, life insurance companies, private pension funds, state and local government pension funds, federal government pension funds, mutual funds, etc., as shown in the Federal Reserve's "U.S. Financial Accounts" Table L. 213 from 1985 to 2020. The cumulative net issuance of US stocks is represented by the sum of net issuance in Table F.213 of the Federal Reserve's US Financial Accounts from 1985 to 2020.

Considering other possible factors that might be able to systematically influence DJIA prices, we found that factors such as economic growth, the fundamental basis of companies, and investor sentiment can have a significant impact on long-term price trends. While some of those economic variables cannot be directly observed, such as investor sentiment, in accordance with the original variables, we have found several strong proxy metrics that can, at least, mainly characterize changes in those macroeconomic variables. This article introduced the United States Gross Domestic Product (GDP), Producer Price Index (PPI), and Investor Sentiment Index (ISI) as control variables representing economic growth, corporate fundamentals, and investor sentiment, respectively, into the model, of which all, GDP and PPI reflecting the inflation parts that caused the nominal increase, and by controlling the ISI factor, fluctuations can then be controlled, leaving our main explanatory variables with a purer causal relationship that directly lead to an upward trend by demand and supply. Please note that data for all of the control variables are obtained from the Wind and Choice databases.

The symbolic representation, definition, and algorithm of specific variables are shown in Table 3.

**Table 3** Symbolic representation, meaning, and explanation of regression variables

| **Type** | **Name** | **Notation** | **Further Explanation** |
|---|---|---|---|
| Dependent Variable | Dow Jones Industrial Average | DJIA | Representing the price of main US stocks |
| Main Explanatory Variable | Net Growth in Total Demand | NID | Representing continuous growth in total demand |
| | Accumulated Net Issuance | CNI | Representing a continuous decrease in the total supply |
| Controlled Variables | Gross National Product | GDP | Nominal GDP by the expenditure method |
| | Producer Price Index | PPI | Producer Price Index without inflation correction |
| | Investor Sentiment Index | ISI | The higher the willingness to invest the higher |

…

Since what we are studying is the causal effects of stock prices and the greatest indicator is a comprehensive stock index, such as the DJIA index, that absorbs all pricing factors in the market, hence, the stock index can be fundamentally regarded as an equilibrium, average price of all tradable stocks in the market. In the supply and demand model, the market equilibrium price is determined by strengths of supply and demand, so the stock price index is essentially generated at the equilibrium of $S_{Stock}(P) = D_{Stock}(P)$. Since total demand is the sum of the demands of all investment entities, and total supply is the sum of the circulating stocks of all listed companies, the two should be equal to the total market value of the stocks. Therefore, the two should be equaled to an identity equation:

$$D_{Stock}(P) = \sum Demand_i\,(P) = S_{Stock}(P) = \sum Supply_i\,(P) = TMV$$

Of them, TMV (Total Market Value) represents the sum of the market value that each company has. The rationale for obtaining equilibrium prices by simultaneous supply and demand implies that the factors that affect the total demand for a stock and the factors that affect the total supply of a stock can naturally be jointly regressed. Therefore, a multiple linear regression model (**Model I**) simply including supply and demand factors is constructed to study the impact of continuous growth in total demand and continuous reduction in total supply on the continuous growth of the US stock index in the US stock market, which is shown below:

$$DJIA = \beta_0 + \beta_1 NID + \beta_2 CNI + \beta_3 GDP + \beta_4 PPI + \beta_5 ISI + \mu \qquad \textbf{Model I}$$

Of those notations, $DJIA$ is the dependent variable, $NID$ and $CNI$ are the explanatory variables, $GDP$, $PPI$, and $ISI$ are regarded as controlled variables, $\beta_0$, $\beta_1$, $\beta_2$, $\beta_3$, $\beta_4, \beta_5$ are coefficients, and $\mu$ is a random term.

…