Non-Value-Added Tax to Improve Market Fairness
Iryna Veryzhenko, Arthur Jonath, Etienne Harb

To cite this version:
Iryna Veryzhenko, Arthur Jonath, Etienne Harb. Non-Value-Added Tax to Improve Market Fairness. 2020. hal-02881064

HAL Id: hal-02881064
https://hal-cnam.archives-ouvertes.fr/hal-02881064
Preprint submitted on 25 Jun 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers. L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.
Non-Value-Added Tax to Improve Market Fairness*

Iryna Veryzhenko†  Arthur Jonath‡  Etienne G. Harb§

Abstract

Promotion of both market fairness and efficiency has long been a goal of securities market regulators worldwide. While previous studies mostly focused on market efficiency, our paper proposes tools to improve market fairness. We define market fairness as the ability of a market structure and its regulatory framework to guarantee unimpeded competition while curbing excessive speculation and market manipulation. Such behaviors undermine the quality of financial markets in the sense that they cause volatility and lead to instability. To encourage value generation and improve market quality, we advance a graduated Non-Value-Added Tax. The proposed tax is implemented in a simulation-based model whereby a profitable transaction is taxed at the higher rate if it does not enhance efficiency measured by deviation from fundamentals. When an agent locks in profit not supported by fundamentals but driven by trend-following strategies, the generated profit is taxed at graduate rates under the Non-Value-Added Tax regime. Unlike existing Financial Transaction Taxes, the Non-Value-Added Tax is levied on profit and not on price. More concretely, our findings show that this tool encourages profitable trades that add-value to the market and discourages valueless profit making. It significantly curtails volatility, and prevents the occurrence of extreme market events like bubbles and crashes.

Keywords: market fairness, market regulation, Non-Value-Added tax, high-frequency trading, bubbles and crashes, efficiency.

JEL classification: G14;D84;D85;E47;I31

*Concepts instrumental in developing Non-Value-Added Taxation as a counter to economic instabilities resulted from a study to incorporate profit motive and human desires into economic theory. We wish to thank George Akerlof and Robert Shiller for their book "Animal Spirits", in which we find support for our approach. The following people helped by reviewing our concepts and their connection to NVAT. Richard Goldwater, M.D. Psychiatry and Co-Founder of Profit and Entropy, explained how the human mind distinguishes between subjectivity and objectivity. Donald Weeden, Chairman Emeritus Weeden and Co. and pioneer in Silicon Valley financing, offered wisdom on role of finance in providing liquidity, value, and worth. Theodore Lehman, former CEO of T.H. Lehman and Co., encouraged adding psychological basis for post-Enlightenment thinking. David Golden, Consulting Professor, Mechanical Engineering, Stanford University, challenged how we brought subjectivity into objective science. Robert Wolf, Prof. Emeritus, Physics, Harvey Mudd College, probed our approach connecting cash flow to heat flow. Michael Reich, Professor, Economics, U.C. Berkeley, asked the questions that led to improved statistical testing of our model. The late Ab Kader, energy-trading software engineer during the Enron debacle, introduced us to paradoxes in trading mentality. The late Nobel Laureate Kenneth Arrow, challenged us to better distinguish between value-added and non-value-added profits. We also wish to thank William Barnett and Fredj Jawadi. We are grateful to George Tauchen for fruitful discussions at the "Financial Markets and Nonlinear Dynamics" (FMND), which greatly improved our model and results.

†iryna.veryzhenko@lecnam.net, Corresponding author, Conservatoire National des Arts et Métiers, LIRSA, Labex ReFi, Paris, France

‡aj@profitandentropy.com, Profit and Entropy, Portola Valley, CA, USA

§egerbanharb@ndu.edu.lb, Notre Dame University, Lebanon
1 Introduction

The investigation of "High Frequency market microstructure" sheds light on two policy issues of particular interest: market linkages and market fairness (O’Hara, 2015). When markets become faster, they don’t necessarily become fairer. While the existing literature focuses on market quality, scarce are the studies that address market fairness, the latter being difficult to apprehend and measure.

Frankfurter (2006) presents the Theory of Fair Markets (TFM) as an alternative to Efficient Market Hypothesis (Fama, 1970). The author argues that the statements "let the market alone" and "market knows better" "are a myth created and nurtured by those who want to take advantage of their political power to keep any regulating body off their hands". To promote market fairness, a regulatory framework should be established to guarantee fair and unimpeded competition, which can improve the allocation of resources and eliminate the opportunistic trades. In light of the global meltdown in 2008, several studies such as Mullineux (2010) and Marti and Scherer (2016) discussed the need for financial market regulations to protect individuals and businesses against the "monopoly powers of large suppliers" and the complexity of some structured financial products. Since then, the connection between financial markets and social welfare gained traction as many researchers such as (Jonath and Goldwater, 2018) devoted close attention to study new regulating mechanisms to combat financial instability and prevent future financial crises, and more particularly promote financial market transactions that add value to society. Marti and Scherer (2016) stress the importance of linkages between the financial innovation (including new types of derivatives, new processes, new market organization and new regulations) and social welfare.

In July 2015, Mary Jo White who served as the 31st Chair of the SEC, called for market reforms to curb unfair advantages of HFT, including reining in high-frequency trading itself and monitoring dark pools and other prohibited trading practices in the world’s largest stock market. Yet, she clearly stated that "the SEC should not roll back the technology clock or prohibit algorithmic trading, but should assess the extent to which computer-driven trading may be working against investors rather than for them". A recent publication by Ladley (2020), reveals that existing regulations such as financial transaction taxes, make/take fees, minimum rest time, increasing the tick size are "ineffective in either improving social welfare, market quality or reducing the profits of HFTs". Our paper contributes to discussions on such non-trivial issues like market fairness and social welfare in financial markets where both concepts seem to be compromised. We develop a tool, the Non-Value-Added Tax (NVAT) inspired by the work of Jonath and Goldwater (2018), to reduce the negative effects of High Frequency Traders, i.e. market instability and market manipulations, without losing the benefits these agents bring to the market, namely market efficiency and price discovery.

It is obvious then to start by defining the term "Value" in our model. In fact, "Value" in financial markets refers to trades that improve market quality; in other terms, trades that contribute to informational efficiency and price discovery, provide for market liquidity, and reduce market volatility. The NVAT we are developing to curb speculative activities is quite different from the Financial Transaction Taxes (FTT) widely discussed and implemented, most notably in the European Union (EU). The main difference is that the NVAT is levied on profit, not on price, which assimilates it to an income tax rather than a sales tax. Defined as a tax on profit, it doesn’t suffer from the main drawback of FTTs which application may result in paying taxes even when incurring losses. Additionally, the NVAT simplifies reporting and transparency requirements, reduces administration costs, and min-
imizes harmonization complexities.

Several articles focused on financial transaction taxes that were applied on prices, but not on profits. Their impact on the volatility and the market liquidity show mixed results. With regard to volatility, Dooley (1996), Kupiec (1996), Subrahmanyam (1998), Amihud and Mendelson (2003) and many others identify a negative effect of FTT implementation on the market liquidity, automatically amplifying the market volatility by driving away rational agents. The latter is supported by Baltagi et al. (2006), Pomeranets and Weaver (2011) and Huber et al. (2014) who correlate volatility with transaction taxes. On the other hand, few articles such as Roll (1989), Saporta and Kan (1997), Liu and Zhu (2009), relate an inverse or insignificant relationship between financial transaction taxes and volatility, after investigating different markets and locations.

As for the impact of FTT on market liquidity and informational efficiency, literature is relatively scarce. We name Frino and West (2003), Bloomfield and Wang (2006), Baltagi et al. (2006), Liu and Zhu (2009), and Pomeranets and Weaver (2011) who find a negative impact of FTT/transaction costs on the bid-ask spread, as a measure of market liquidity as well as on the informational efficiency. Veryzhenko et al. (2017) implemented a tax on canceled orders by High Frequency Traders (HFTs), using an agent-based financial modeling. The authors show that HFT liquidity is short-lived and that the implementation of tax reduces HFT activities, which seems to have an insignificant impact on market volatility and market liquidity as measured by bid/ask spreads, while only dollar volumes decrease. The authors also show that reducing HFT activities lead to less efficient markets as the deviation from fundamentals increases.

Most of these FTTs discussed in the above literature are variations of the currency taxes such as Spahn (1995) and Tobin (1978), designed during the past several decades to keep in check the excesses in the currency market. While a Tobin transaction tax may reduce market liquidity, it also limits the desired stabilizing effect of such a tax. Demary (2011) confirms the observations of Westerhoff and Dieci (2006) that under a tax rate of 0.1%, about 80% of foreign exchange traders forego trading, thus decreasing liquidity on the foreign exchange market. Finally, in 2015, the Tax Policy Center of the Urban Institute and the Brookings Institution published a summary of the general theory behind the FTT and described its status among the G-20 and major economies worldwide as well as its revenue potential in the U.S. (Burman et al., 2015). The report also offered a comparative study of the implementation methods in use and of those, which were still under discussion. It shed light on the main weaknesses of the FTT approaches due to their rates and their application to gross rather than net revenue. The author concludes that the FTT, at the rates being proposed, would discourage all trading, not only speculation and rent seeking; would increase market volatility rather than curbing it; and would possibly create new distortions among asset classes and across industries, thus hurting market fairness and social welfare. The findings of Ladley (2020) also confirm that FTT reduces the profits of all categories of traders, high-frequency and slow traders.

Finally, according to Burman et al. (2015), the FTT appears to be poorly targeted at the financial sector excesses that led to the Great Recession. That said, it was suggested to replace it by a Financial Activities Tax (FAT) or Value Added Tax (VAT), assuming that these taxes might be more effective and less distortionary1, if the goal of the tax is solely to have the financial sector pay and compensate the rest of the economy for the costs that

---

arose from the financial crisis.

Both the FAT, defined as a profit-associated tax on the sum of bankers’ excessive remuneration and bank profits, and the VAT, defined as a price-associated tax applied to the sum of profit and costs, provide no distinction between profits on transactions generating value and those that which do not. Also, in taxing both profit (capital income) and costs (labor value) without differentiation, VAT adds financial incentive to cut costs. In such a way, the VAT described as a regressive sales tax, encourages the forces that clearly act to increase the capital-to-labor wealth gap. NVAT, a more practical FTT, optimistically could mitigate market instabilities such as those caused by rapid expansion of non-value-added profit bubbles.

In light of the above limitations of existing tax regimes and their ineffectiveness as FTTs in improving market quality, the NVAT appears to be superior with regard to targeting a fairer and enhanced financial market quality as well as encouraging profitable financial transactions that add value to society. Our model investigates implementing the NVAT, which is also meant to limit and help meet the costs of future crises, similarly to some traditional FTTs, as called for by the G-20 ministers in April 2010. We show that NVAT tax brackets are tied to a ratio of cost to profit, both of which are recorded for each transaction. No new bureaucracy is required for data acquisition or transparency.

Clearly, the NVAT addresses two major drawbacks forcing some governments to consider withdrawal from the FTT they implemented. First, the NVAT removes the concern about rate variations applied to different financial instruments such as stocks vs. derivatives. With NVAT’s profit-based tax rate table, divided into tiers of value-added content in transactions, the identity of specific financial instruments does not matter. Second, it lessens the confusion and difficulty of harmonizing the tax base among transacting institutions, in different territorial locations. Since the tax is applied to profit, the NVAT tax only applies to profitable transactions (sell associated with initial buy, and buy associated with initial short selling). Buyer location is of no concern.

Since NVAT applies to profitable trades, let’s clarify how our model handles a trade’s profit. We compare the profit of a trade to the variation in fundamental value between the two points of the trade. If the price variation is higher than the variation in fundamentals, the agent will realize an extra profit, not supported by fundamental information. This extra profit can be explained by momentum or a trend component. The momentum-based extra profit is definitely taxed based on NVAT tax rate regimes.

In this paper, we propose an agent-based model to analyze the effects of introducing the NVAT tax on stock markets, from both structural and behavioral perspectives. We offer to contribute to the FTT debate by studying whether and how market volatility and trading activity are influenced by the suggested profit-based tax. We compute a wide range of volatility and efficiency measures to grasp different dimensions of market quality. We develop a simulator that acts as an artificial financial market. This computational-experimental approach enables us to perform several validation tests and some hypotheses in order to give insights to regulators with regard to fiscal regulatory policies. Noticeably, the tax we examine deals better with the objectives of stabilizing the market, discouraging speculation and improving market fairness.

In the following sections, we first set representative models for implementing NVAT as a financial transaction tax applied to traders’ profits. Second, we discuss the results of simulations in terms of tax collecting capacity and the impact of the tax on market quality.

---

2 Ibid. pp. 8,9
We also examine its effect on market quality in extreme market events like flash crashes and bubbles. Finally, we draw conclusions on how the NVAT can be applied to influence trading behaviors, in favor of an increased market fairness and a reduced income inequality gap that is growing more and more worrisome.

2 Experimental design

We run a series of experiments to capture the effect of the NVAT tax on traders’ profits and market quality. First, we consider a set of simulations where we ignore taxes. This set serves as a benchmark. Then, we incorporate taxes and run the simulations under the same initial settings and compare the results and observations with the benchmark, i.e. the initial untaxed market in order to examine the effect of profit-based tax on market quality.

We develop a simple synthetic model of the market, with only one risky security and \( N \) competing traders subject to the same institutional market design. The trading agents belong to a heterogeneous population and participate within a protocol organized in trading sessions or rounds. In each session, traders exchange only one unit of equity.

2.1 Market mechanism

In this paper, we use the ArTificial Open Market platform (ATOM)\(^4\) introduced by Brandouy et al. (2013). Developed as a large scale experimental platform, ATOM offers three main interacting modules: (i) the market microstructure, whereby we define the mechanism of order routing and price fixing, (ii) the economic environment that generates exogenous information on corporate developments, dividend payout policy, and coupon changes, and (iii) an agent component, which offers multiple types of agents with different utility functions, views, and strategies. Traders react to exogenous information (such as expected returns, estimated risks, etc.) and endogenous information (such as post transaction information generated by agents’ interactions), imposed restrictions, and market mechanism rules.

Central order book

A market mechanism comprises the set of rules that defines how we transform agents’ orders into series of transactions with their timing, price and volume. While in most of simulation-based studies (Lux and Marchesi, 2000; Pellizzari and Westerhoff, 2009; Pouget, 2007) clearing price is computed based on aggregate excess demand and supply, we reproduce in our paper, a realistic central order book market mechanism. The Central order book we develop represents a continuous trading mechanism when multiple transactions are possible at each time step. All orders, called limit orders, are stored for execution according to a strict price-time priority on the two sides of the order book, i.e. the Bid (demand) and Ask (offer). The highest demand price represents the best bid and the lowest offer price represents the best ask. When a new buy or sell order arrives to the market, the execution conditions are checked. The market price is updated continuously and all previous and current orders, timing, volume and price of transactions are always visible to the public.

Time scale

A key element in all multi-agents systems is the scheduler which models the time scale. It manages the very moment when agents act, orders are executed, and the price is fixed. A scheduler can be treated as a set of loops (or rounds) in the simulations; in each of them,

\(^4\)http://atom.univ-lille1.fr
agents can express their decisions to buy, sell or do nothing. ATOM is also able to start each round by activating the same category of traders. In this way, we can simulate the privileged access of HFT to the order book.

Moreover, the number of rounds determines the time granularity of simulations. We consider 1,000 time-steps, which approximately correspond to a half-minute time scale or 8.5 hours of trading session. Since the choice of the trading frequency made by agents is essential for them to reach their investment objectives, we allow them in ATOM to decline the suggested trade decision generated by the simulator. In such a manner, they can choose and set their trading frequency. In our simulations, the trading frequency of fundamentalists (described below) varies from once per minute to once per hour while the trading frequency for trend followers is set at a half-minute, as this type of agents can trade at the finest time grain.

2.2 Traders behavior

In the stock market, the main role of agents is to analyze the information they get and make decisions in line with their selected strategies to translate their knowledge into buy and sell actions. Agents gather information and react relatively sensibly to it. The price dynamic is then a result of non-trivial interactions between traders, market microstructure and regulatory rules. Furthermore, the price dynamic becomes itself a source of information (momentum signal) for investors, who trade accordingly and end up affecting the market dynamic. Consequently, this creates a feedback loop.

Additionally, another type of information can potentially guide the traders’ decision making; it is the firm-specific or fundamental signal. We model a fundamental value dynamic as follows $V_{t+1} = V_t + \sigma_t$, $\sigma_t \sim U[-1;1]$, for $t = 1, \ldots, N$, where $V_t > 0$ and the initial fundamental value is $V_0 = 200$.

Fundamental value is organized in a matrix $1,000 \times 1,000$ (for 1,000 rounds and 1,000 repetitions) in all scenarios. Traders can trade based on this information or completely ignore it, depending on their heterogeneous preferences.

**Fundamentalists** are motivated by the real (fundamental) asset value. The fundamental value of a stock follows a jump process $V_{t+1} = V_t + \delta_t$, where $\delta_t \sim N(0, \sigma^\delta)$ is a normal random variable with zero mean and constant standard deviation. A $\delta_t > 0$ signals a positive prospect, thus investors expect a price increase. A $\delta_t < 0$ signals a negative prospect and a price decrease. The case of $\delta_t = 0$ denotes an ambiguous message that ends up being ignored by investors when they form their own expectations.

Since we assume that agents are boundedly rational (or noisily informed), the fundamental value appears to be biased by $\epsilon_i$, which determines the accuracy with which the agent $i$ interprets the fundamental information $E_{ij}(P_{t+1}) = V_{t+1} + \epsilon_i$, $\epsilon_i \sim N(0, \sigma^\epsilon)$. Agents belong to a heterogeneous population with respect to their parameter $\epsilon_i$, which is normally distributed with mean zero and constant standard deviation $\sigma^\epsilon$.

To decide whether to be short or long on a stock, an agent compares the current stock price $P_t$ with his expectations $E(P_{t+1})$. If $P_t > E_{ij}(P_{t+1})$, the stock is overvalued, hence the agent can benefit from this deviation from fundamentals by placing a sell (Ask) order. If $P_t < E_{ij}(P_{t+1})$, the stock is undervalued, hence the agent takes advantage of it by placing a buy (Bid) order.

**HFTs** or high-speed trend followers do not take into account the fundamental value, they try instead to detect trends and trade accordingly. They rely on historical price dynamics to anticipate future price variation.
\[ \frac{P_t - P_{t-n}}{P_{t-n}} > \Delta_i. \] (1)

The agents are heterogeneous with respect to the parameter \( \Delta_i \) of the minimal price variation and its interpretation. In our model, trend followers buy (sell) when the stock value has been increasing (decreasing) over the last 10 to 100 rounds, which is reflected in the way we set the parameter \( n \). Their behavior can be described as a positive feedback trading as they detect a trend and they reinforce it. Trend followers are constantly seeking trading opportunities, and define their expectations as follows:

\[ E_{i,t}(P_{t+1}) = P_t + \psi_i(P_t - P_{t-1}) \] (2)

where \( E_{i,t}(P_{t+1}) \) represents the expectation of the agent \( i \) at the moment \( t \) about the future price \( P_{t+1} \). \( P_t \) is the market clearing price at the moment \( t \). \( \psi_i \) is the sensitivity of the agent \( i \) to the market trend \( P_t - P_{t-1} \). This parameter \( \psi_i \) can be also interpreted as the importance given to the momentum signal. This parameter is a central one in our model where it identifies the speculation activity. It is uniformly driven from the interval \([0;1]\). Investors sell (buy) when the last past clearing price is higher (lower) than his focal price expectations, \( P_t > E_{i,t}(P_{t+1}) \) \( (P_t < E_{i,t}(P_{t+1})) \). To sum up, traders would be interested in buying undervalued stocks and selling overvalued stocks, based on their beliefs.

In our simulations, the agents apply adaptive order submission expecting to maximize their profit. Yet, traders face a trade-off between the potential profit and the tax rate, as a higher profit means graduating into an upper tax bracket. Hence, they don’t adhere to their expectations. They check the current state of the order book to optimize their final trade. In a double auction market, a profit-oriented buyer would propose a price that does not fully reflect his reservation or expected price (the maximum price he would be willing to pay), betting on the existence of a seller who would accept to fill this relatively low bid order. Similarly, a seller would propose a price that is higher than his reservation or expected price, betting on the existence of a bidder ready to buy at this relatively high price. Agents set the direction and price of their orders based on the last market price and the current state of the order book. This behavior is summarized in Table 1.

### 2.3 Non Value Added Tax (NVAT)

We next incorporate the NVAT into our model, whereby all traders involved in transactions update their cash and stocks’ positions with respect to transaction price and transaction tax (NVAT). For each agent, we compute the profitability of the trade: buy low and sell high (as margin and short selling is not allowed) \( \Delta_P = V_k(P_k - P_l) \) (dollar volume variation) where \( V_k \) is the volume sold at the moment \( k \) at price \( P_k \) and initially bought at price \( P_l \) at the moment \( l \). We compare the profit of the trade to the variation in the fundamental value between these two points of time \( \Delta_F = V_k(F_k - F_l) \). If \( \Delta_P > \Delta_F \) it means that an agent was able to realize an extra profit not supported by fundamental information. The latter can be explained by momentum or trend component of the transaction. For instance, upward trend is possibly initiated by the positive information, specific to the concerned company (microeconomic level information), which can be exploited (confirmed and reinforced) by the trend followers (the technical analysts). We can then tax the momentum-based extra profit, according to NVAT tax regimes. Three of these regimes are proposed in Table 5.

---

5 As we reproduce central limit order book market mechanism, agents’ heterogeneity is a key element to guarantee a continuous trading.
Table 1: The order-submission procedure. \( P_D \) denotes the best bid price, \( P_S \) is the best ask price, \( P_D \) is a set of bid orders (demand), \( P_S \) is a set of ask orders (supply), \( \emptyset \) denotes empty set, \( U(x_1, x_2) \) is the uniform distribution in the interval \([x_1, x_2]\), and \( E_{i,t}(p_{t+1}) \) represents the expectation at moment \( t \) of the agent \( i \) about the future price \( p_{t+1} \).

| Table 2: Examples of NVAT tax rate, where \( \%F \) is the variation in fundamentals between two trades, \( \%P \) is the percentage price variation between two profitable trades. |

2. In our simulations, the NVAT is applied to the total profit, that is, NVAT payment = \( V_k(P_k - P_l) \times \text{NVAT rate}. \) Each of these experiments consists of 1,000 runs, each starting with the same initial conditions (initial wealth, stocks held, agent population) except the fundamental value. Hence, we average all the statistics of 1,000 repetitions. The parameters we use in our simulations are listed in Table 3. The parameters’ estimation is in line with empirical papers such as (AMF, 2017; Colliard and Hoffman, 2017; Kirilenko et al., 2017) and those using agent-based modeling research such as (Oriol and Veryzhenko, 2019; Pellizzari and Westerhoff, 2009; Veryzhenko et al., 2017).

3 Results and discussion

As we previously mentioned, we focus on three dimensions of market quality: market volatility, market efficiency and market liquidity. For each experimental set, we generate 1,000 series of runs per day. Based on these series, we compute multiple metrics of market
<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$N^{FD}$</td>
<td>1,000</td>
<td>Number of fundamentalists</td>
</tr>
<tr>
<td>$N^{HFT}$</td>
<td>200</td>
<td>Number of HFTs</td>
</tr>
<tr>
<td>$C_{0,i}$</td>
<td>20,000</td>
<td>Initial cash attributed at moment 0 to agent $i$</td>
</tr>
<tr>
<td>$S_{0,i}$</td>
<td>100</td>
<td>Number of stocks attributed at moment 0 to agent $i$</td>
</tr>
<tr>
<td>$N_{rounds}$</td>
<td>1,000</td>
<td>Number of rounds per day</td>
</tr>
<tr>
<td>$V_0$</td>
<td>200.00</td>
<td>Initial fundamental value</td>
</tr>
<tr>
<td>$\varepsilon_i$</td>
<td>[-1;1]</td>
<td>Accuracy of fundamental value prediction by the agent $i$</td>
</tr>
<tr>
<td>$\Delta_i$</td>
<td>[0,0.001]</td>
<td>HF traders’ activation threshold</td>
</tr>
</tbody>
</table>

Table 3: Parameters and their initial values used in simulations

As a proxy for volatility, we use the average of absolute returns and squared returns across each trading period:

$$|R_k| = \frac{\sum_{t=1}^{T} |R_{t,k}|}{T}$$

and

$$R_k^2 = \frac{\sum_{t=1}^{T} R_{t,k}^2}{T}$$

where $t$ denotes each transaction, and $T$ measures the total number of transactions within a given period $k$, $|R_{t,k}| = |\ln(P_{t,k}) - \ln(P_{t-1,k})|$ and $R_{t,k}^2 = (\ln(P_{t,k}) - \ln(P_{t-1,k}))^2$. Another volatility measure commonly used in the literature (LiCalzi and Pellizzari, 2007) is the standard deviation of returns over the given period:

$$\sigma_k = \frac{\sum_{t=1}^{T} (R_t - \bar{R})^2}{T-1}$$

Similarly to LiCalzi and Pellizzari (2007), we assume that $k$ represents the returns over the last 20 rounds. For comparability’s sake, all data is aggregated (averaged) at a half-second grain, where 20 rounds represent a 10-minute time window. It determines the size of the moving window we compute to create a series of statistics.

We measure informational efficiency by the absolute deviation between the price $P_{t,k}$ and the fundamental value $F_{t,k}$

$$\frac{1}{T} \sum_{t=1}^{T} \left| \frac{P_{t,k} - F_{t,k}}{F_{t,k}} \right|$$

Finally, market liquidity is computed as the total daily trading volume.

**Difference-in-differences.** In order to isolate the changes in metrics due to the policy implementation, we use the Difference in Differences (DiD) technique (Ashenfelter and Card, 1985). The advantage of DiD is that instead of comparing of averages of units, it compares the differences in means of units of treated and control groups over time. We consider two groups and a 2-period case. Only one population is subject to the NVAT:

$$Y_i = \beta_0 + \beta_1 \cdot D^{treated} + \beta_2 \cdot D^{tax} + \tau \cdot D^{treated} \cdot D^{tax} + \varepsilon_i$$

We regress the metrics of efficiency, volatility and liquidity $Y_i$ on a set of treatment indicators, which include a dummy variable identifying the treated group $D^{treated} \in \{0, 1\}$, a dummy indicating an after tax period $D^{tax} \in \{0, 1\}$ and the interaction of these two dummies $D^{treated} \times D^{tax}$, $\tau$ being the parameter of interest. If the tax has a significant effect on the dependent variable, the regression will return a significant coefficient of the $treated \times tax$. 

9
Table 4: The impact of the introduction of different tax regimes on stock market efficiency, volatility and liquidity. Difference in differences analysis. (s.e.) = standard error of test. $R_t = \log(P_t/P_{t-1})$ is the log return. $\sqrt{\text{return}}$ and $\text{Abs. return}$ are respectively the squared and absolute returns. $\sigma_k$ is the standard deviation computed based on 20 returns with a moving window of $k = 20$. The deviation from fundamentals represents the absolute difference between the fixed price and the fundamental value at the moment $t$. $\text{Volume}$ represents the daily trading volume. Column $\text{Tax}$ represents the effect of the VAT on metrics of market quality. (s.e.), adj. $R^2$ and $p-value$ show respectively the standard error, the adjusted $R^2$ and the p-value of the test.

<table>
<thead>
<tr>
<th>Regime 1</th>
<th>Regime 2</th>
<th>Regime 3</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Abs. return</strong></td>
<td><strong>Abs. return</strong></td>
<td><strong>Abs. return</strong></td>
</tr>
<tr>
<td>Tax (s.e.) 4.991e-05</td>
<td>Tax (s.e.) -9.077e-05</td>
<td>Tax (s.e.) -9.101e-05</td>
</tr>
<tr>
<td>$p-value &lt;2e-16$</td>
<td>$p-value &lt;2e-16$</td>
<td>$p-value &lt;2e-16$</td>
</tr>
</tbody>
</table>

| **Abs. return** | **Abs. return** | **Abs. return** |
| Sqrt return | Sqrt return | Sqrt return |
| Tax (s.e.) -3.369e-07 | Tax (s.e.) -5.997e-07 | Tax (s.e.) -8.258e-07 |
| $p-value <2e-16$ | $p-value <2e-16$ | $p-value <2e-16$ |

| **Abs. return** | **Abs. return** | **Abs. return** |
| **Deviation from fundamental, %** | **Deviation from fundamental, %** | **Deviation from fundamental, %** |
| $\sigma_k$ | $\sigma_k$ | $\sigma_k$ |
| Tax (s.e.) -4.829e-05 | Tax (s.e.) -8.083e-05 | Tax (s.e.) -8.258e-05 |
| adj. $R^2$ 7.341e-06 | adj. $R^2$ 7.374e-06 | adj. $R^2$ 7.442e-06 |
| $p-value <2e-16$ | $p-value <2e-16$ | $p-value <2e-16$ |

| **Abs. return** | **Abs. return** | **Abs. return** |
| **Volume** | **Volume** | **Volume** |
| Deviation from fundamental, % | Deviation from fundamental, % | Deviation from fundamental, % |
| Tax (s.e.) 0.073625 | Tax (s.e.) 0.09409 | Tax (s.e.) 0.07187 |
| adj. $R^2$ 0.013028 | adj. $R^2$ 0.01228 | adj. $R^2$ 0.01478 |
| $p-value <2e-16$ | $p-value <2e-16$ | $p-value <2e-16$ |

| **Abs. return** | **Abs. return** | **Abs. return** |
| **Volume** | **Volume** | **Volume** |
| Tax (s.e.) -387.7 | Tax (s.e.) -695.3 | Tax (s.e.) -826.94 |
| adj. $R^2$ 124.9 | adj. $R^2$ 120.2 | adj. $R^2$ 135.46 |
| $p-value <2e-16$ | $p-value <2e-16$ | $p-value <2e-16$ |

$\text{Volume}$ represents the daily trading volume. Column $\text{Tax}$ represents the effect of the NVAT on metrics of market quality. (s.e.), adj. $R^2$ and $p-value$ show respectively the standard error, the adjusted $R^2$ and the p-value of the test.

Signif. codes: 0 ‘***’, 0.001 ‘**’, 0.01 ‘*’, 0.05 ‘.’, 0.1 ‘ ’, 1
Based on Figures 1(a) - 1(d) reported in the Appendix, we identify a positive effect of NVAT on controlling volatility (in all the 3 tax regimes). In fact, results of Difference-in-Differences test reported in Table 4 show that the absolute return and the squared return decrease by 2.4% and 2% respectively with the implementation of NVAT. The standard deviation also decreases by 1% on average. However, it negatively influences market liquidity, reducing the trading volume. Hence, based on all quantitative metrics, we conclude that market volatility significantly decreases with the introduction of an additional tax. This can be explained by a reduced trading activity. NVAT has a direct effect on speculators, dissuading them from entering into unproductive trades. We also identify a reduction in the volume of orders and in the trading frequency. According to DiD results, the trading volume decreases on average, by 2.3%. Yet, the decreased trading volume should not be perceived as a simple dry-up of liquidity. The countervailing side is that a reduction in volume of exchange means less frequent order submissions, updates, and cancellation of orders that may lead to dangerous market fluctuations. The graduated NVAT tax rate tables provide the tools for marketplace supervision to manage balancing tradeoffs between liquidity and frenzy. In this manner, NVAT removes some part of speculative non-productive volume. The reduced speculative activity leads to a calmer and less volatile market, as a direct effect of NVAT introduction. The statistics in Figures 2 and Figures 3 do not show a significant improvement of market efficiency. The quantitative results of DiD test in Table 4 reveal that the introduction of NVAT appears to have a rather negative effect on market efficiency, increasing the deviation from fundamentals. Furthermore, NVAT impacts not only speculators but it also dissuades some fundamentalists (who sell overvalued and buy undervalued stocks, based on their beliefs) from executing some usual trades. As a result, the deviation from fundamentals increases, thus amplifying the NVAT collection.

**NVAT and Profitability of Trading.** To understand the effect of the NVAT on profitability of different strategies, we measure the end of period return, computed as \( \frac{W_{i,T} - W_{i,0}}{W_{i,0}} \), where \( W_{i,0} \) is the initial wealth of agent \( i \) and, \( W_{i,T} \) is the end-of-day wealth of agent \( i \). Figure 4 reports the distribution of end-of-period realized returns to Fundamentalists and High Frequency Traders. Figures 4(a) - 4(b) illustrate the effect of NVAT on profitability of trading. We do not find a significant difference in means for both categories of traders. The average of the end-of-period return of HFTs in taxed market is -0.46%, in untaxed market -0.55%. The p-value of t-test, used to measure the equality of means of two series, is of 0.694. However, the results show that the NVAT is able to prevent traders from extreme losses.

**How much tax revenues would the NVAT allow the government to collect?** One of the intentions of any tax is to raise substantial revenue. In order to estimate how much money the NVAT would collect, we measure an average daily tax payment made by different categories of traders. The findings are summarized in Figure 5. Note, that these figures are computed based on our artificial market model and can be scaled to markets of different sizes. On average our representative group of high frequency traders pays 45,000 US dollars with $8,000 standard deviation, while the group of slow fundamentalists pays $380 with $151 standard deviation. Fundamentalists and HFTs pay annually, an average of $132,805 and $15,488,106, respectively. Figures 5(a) - 5(b) illustrate a distribution of tax revenues collected from different categories of traders.

The results in the tables show that the NVAT significantly increases tax revenues, without causing major distortions. Moreover, NVAT clearly improves market volatility.
4 Extreme price movements

In the present section, we examine the effect of the NV AT on market quality, in extreme market events like flash crashes and bubbles. In accordance with Bellia et al. (2018), we identify a mini-bubble (crash) as a strong and rapid price increase (drop), at least by 1.5% of the initial level, followed by a violent burst (recovery), within at most 12 min (24 rounds in our simulations). To be identified as a bubble (crash), the log-price should retrace at least one-third of its initial rise (decline) within the above mentioned time window. Such extreme price movements may spring from practices that destroy liquidity. In this section, we consider two sources of potential bubbles and crashes: the first one is defined as an operational error and so called "fat finger", it stems from the submission of a big volume order, which destroys liquidity. The second source of potential market instability is described as agents’ synchronization or mimicking behavior, whereby crashes or bubbles endogenously emerge.

4.1 Exogenous liquidity shock

First, we create an extreme market event by introducing an aggressive market order, similarly to Brewer et al. (2013). Such an order can be a market order with a volume bigger than that available at the best price (Degryse et al., 2005). This order triggers a flow of transactions and causes the price to shift.

In our simulations, we specifically define the bubble as the result of placing a buy market order with a volume 20-times bigger than the average order size. According to Degryse et al. (2005), liquidity measures take about twenty best limit updates to get back to their initial level. This explains the choice of the size or volume of the aggressive market order we need to place to create an extreme event. Such an order produces an immediate effect on the market dynamics. We then assess the reaction of the market prior to and after incorporating the NV AT, and we analyze its ability to reduce market volatility and hinder extreme market movements. We note that in a market not subject to regulations or taxes, the average price increase during a typical bubble, is of 4% with a standard deviation of 1.45%. However, in the taxed market under NV AT regimes, the price mean is 2.02% with a standard deviation of 1%. Hence, we infer that HFTs trend followers optimize their trading decisions and evaluate their potential final profit, under different tax regimes, based on their profit size. They consistently readjust the activation parameter Δ_i with respect to different tax rates. They tend to sell before they move up a tax bracket, even though the price is still increasing. Thus, they execute trades earlier into opposite direction of the underlying trend and ease market correction. Figure 6 plots the median cumulative return, across 1,000 simulations, as a function of time in the harmonized time units, together with 5% and 95% quartiles. The figure shows a steep upward trend, followed by a partial correction.

Consequently, the traders face a trade-off between the profitability of a trade and tax rate. A higher trade profit is subject to a higher tax rate. Under the first tax regime we implemented, the most frequent tax rate applied to the profitable transactions is 25%, which seems to be a potential solution to the trade-off described above. We reveal a significant increase in the daily tax revenues from HFTs up to $56,000 (compared with initial 45,000 US dollars under normal market conditions). The tax seems to dissuade the HFTs from placing 20 million additional orders. We then analyze the net positions of different categories of traders to understand their contribution to the price correction. To measure the liquidity provision and liquidity consumption by different categories of traders, we use monetary
net trade imbalance. It is the difference between the funds invested to buy transactions and funds gained as a result of sell transactions. Negative net imbalance of a trading category, during a bubble, indicates that it contributes to price correction; however, a positive net imbalance, during a bubble, indicates that this category of market participants trade into the direction of the upward trend. We identify negative net positions for both categories of traders, HFTs and non-HFTs, in both scenarios. In the untaxed benchmark scenario, HFTs sell on average for $5,108,292 while non-HFTs sell for $3,235,573. In the taxed market under NVAT regimes, HFTs only sell for $1,323,181, while non-HFTs sell for $2,679,055. Hence, the price increase in the untaxed market is higher, making sell transactions particularly profitable, and the volume of sale positions even higher.

4.2 Endogenous liquidity dislocation

We next focus on the mimicking behavior of agents that can lead to self-reinforced bubbles and crashes (Bookstaber, 2017; Kirman, 1991; Lux, 1995). Similarly to Bookstaber (2017); Brock and Hommes (1998); Lux and Marchesi (2000), we implement the model of contagion driven by the mimicking behavior of traders with regard to expectations: traders can be influenced by trading decisions and performance of other market players, thus they are capable of adopting a more profitable strategy to a certain extent. This switching is driven by the high profits of the mimicked strategies. Hence, the fractions of the resulting strategies vary immediately with the portfolio growth of agents adopting the successful mimicked ones. We assume that agents have a short-term memory and only pay attention to the results of their neighbors at the last round.

Similarly to Watts and Strogatz (1998), we assume a random network to describe the connections between agents. Linkage between agents is defined as a binary matrix $N \times N$

$$M_{i,j} = \begin{cases} 1, & \text{if agents are connected, } \forall i, j = 1, N \\ 0, & \text{otherwise} \end{cases} \quad (7)$$

Like Watts and Strogatz (1998), we assume that each agent is connected with $N/2$ neighbors (where $N = 1,000$ is the number of agents in the simulations) and half of them initially follow fundamentalist strategy while the other half follow trend based strategy. We also assume that the graph structure is not rewired at each time step. Unlike Watts and Strogatz (1998), we assume an asymmetric connection between agents, whereby if agent $i$ can analyze the performance of agent $j$, it is not necessary that agent $j$ can observe the performance of agent $i$.

In line with Lux and Marchesi (2000), we compute the potential profit of the fundamentalist $i$ at moment $t$ as the percentage change in price, which represents for the fundamentalist $i$ the deviation of the last trade price from the expected fundamental value. Since fundamentalists believe that the stock price will converge to its intrinsic or fundamental value, we have:

$$\pi_{FD} = \frac{V_t - p_{i,t}}{p_{i,t}} \quad (8)$$

where $\pi_{FD}$ represents the expected profit of fundamentalist, $V_t$ denotes the expected fundamental value, and $p_{i,t}$ is the price of the pending order at moment $t$, submitted by agent $i$.

As for a trend follower, the profit is computed as the percentage change of his pending order price (or his last trade price) $p_{i,t}$ from last market price $P_i$:

$$\pi_{HFT} = \frac{P_i - p_{i,t}}{p_{i,t}} \quad (9)$$
Consequently, if agent $i$ is initially a fundamentalist at moment $t$, he will compare the past average profit of strategies followed by his neighbors. Like Brock and Hommes (1998); Pellizzari and Westerhoff (2009); Westerhoff (2008), if the average profit of HFTs is higher than that of fundamentalists $\pi_{t-1}^{HFT} > \pi_{t-1}^{FD}$, the agent will switch to trend following with a probability of:

$$\Psi_{t}^{HFT} = \frac{e^{\pi_{t-1}^{HFT}}}{e^{\pi_{t-1}^{HFT}} + e^{\pi_{t-1}^{FD}}}$$

Accordingly, the probability to use fundamental analysis is $\Psi_{t}^{FD} = 1 - \Psi_{t}^{HFT}$.

Figure 7(a) shows the trajectory of the fraction of HFTs over 1,000 time steps in the untaxed market. This trajectory illustrates the strong fluctuations and emergence of clusters of dominant strategies. A joint analysis of Figure 7(a) and Figure 7(b) reveals that with a given proportion of trend followers, the deviation from fundamentals is self-reinforced, which results in mini-crashes or mini-bubbles. Since, the fundamentalist’s potential profit is computed as the percentage change in price deviating from the fundamental value, generating bubbles increases the potential profit to the fundamentalist, who believes that the price will converge to its intrinsic value. The latter encourages the adoption of a fundamentalist strategy, which seems to be temporarily more profitable. Hence, the population of agents converge slowly to fundamental analysts. As a result, the bubble bursts (or the crash is recovered) and the price gets corrected.

In fact, in our simulations, the population of agents never completely converges to one dominant strategy (total contagion of expectations). These results are in line with those of Lux and Marchesi (2000) and the empirical analysis of Cont (2007). In the latter article, the author states that the investor’s inertia explains the switching behavior, which leads to the emergence of clusters of strategies and the persistence of the magnitudes of price changes called volatility clustering. If the volatility is low, agents become more sensitive to news and generate high excess demand, thus increase the amplitude of returns. If the volatility is high, agents become less reactive to news, which increases agents inertia and reduces the amplitude of returns.

That said, what would be the effect of NVAT on the switching behavior (synchronization) and as a result, on the emergence of bubbles and crashes? Our simulations reveal that NVAT is not able to completely eliminate the speculative behavior or to encourage all market participants to adopt long-term fundamental-based strategy. However, the NVAT significantly reduces the profitability of speculators and considerably reduces the size of speculators’ emerging clusters. Consequently, NVAT contributes to regulating the emergence of extreme price events such as bubbles and crashes.

Actually, in the taxed market, HFTs’ population becomes dominant (followed by at least 501 agents) 43.82 times on average, while in the untaxed market, HFTs dominate 59 times on average. If the tax is implemented, the number of HFT agents becomes 209 agents on average, compared to 304 agents in the untaxed market.

Furthermore, in the untaxed market, we determined an average of 34 episodes of extreme price variations of 1.5% (crashes or bubbles) over 12 minutes and 3 episodes with log-price variations higher than 3%. Upon introduction of NVAT, the number of episodes of 1.5% log-price variations is reduced by nearly a factor of 4. The statistics based on 1,000 simulations show that the price declines or increases by more than 1.5% averages 9.27 times, with standard deviation of 10.41 times in the taxed market. In the latter market, we identify only 1.34 episodes of extreme price movements of more than 3% on average.
5 Conclusion

This paper brings attention to the fact that there are "value-added" and "non-value-added" activities, which have quite different effects on our financial system and the economy as a whole. We propose a new financial transaction tax (FTT) to improve fairness in financial markets dominated by short-horizon, profit-oriented traders and to contribute to social welfare by discouraging excessive trading behaviors, which do not add value to the market. We implement the Non-Value-Added Tax (NVAT) in a simulation-based model. NVAT is a graduated tax, whereby a transaction is taxed at a higher rate if it does not enhance market liquidity, preserve market stability, or strengthen market efficiency. We test the effect of the NVAT on order-driven continuous trading, where we show that NVAT significantly reduces the profitability of traders relying heavily on momentum signals, considerably reduces volatility, but slightly decreases the trading volume. Additionally, NVAT reduces the amplitude of extreme market movements resulting from exogenous liquidity shocks like "fat finger" events. NVAT reduces bubble-formation dynamics, and stave off dangerous financial fluctuations from "tipping points". As NVAT is levied on profit and not on price, investors solve an equation of profit/tax rate relationship and not purely maximize their profit. This mechanism makes investors close their positions at the earlier stages of bubbles and contribute to the price correction.

NVAT encourages profitable trades that add value to the market and discourages valueless profit making. It will therefore help provide a common, robust self-regulatory framework that protects investors and improves market fairness. It works as an FTT but levied only on profits, thus NVAT dissipates the rightful concerns held by traders who claim that current FTTs are costly to administer across borders and over a variety of trading products beyond shares. In fact, the application of NVAT only to profits suggests: a) buyers pay no transaction tax until they become sellers, so buyer location has no tax consequence, and b) replacing a tax on price by a tax on profit converts FTT from a sales tax to an income tax. With this in mind, NVAT could significantly contribute to tax revenues without causing major distortions of market quality. The data required for NVAT assessment are recorded at the moment of trades; no new bureaucracy is needed to gather additional or specific data. Policy adaptations may be enacted through adjusting the NVAT rate tables or regimes, which is a common for regulators.

Historically, the mitigation enacted to avoid total financial paralysis endured during the worldwide 2008-9 crash show that governments must step in as last-resort guarantor against catastrophic, non-recoverable, financial institution losses. The truth behind their "Too Big to Fail" description forces the general tax-paying public to act as unwitting partners with financial institutions during their downside calamities. In this respect, NVAT ensures a fairer system, whereby it makes financial institutions also partners with the trading public in its upside profits. This power sharing may exactly fit the need of European Unions’s vision of "a transparent market with clear rights and protections for EU citizens".

References


A Appendix

Figure 1: Evolution of excess volatility. Excess volatility means that it excludes the volatility of fundamental value. Standard deviation is computed based on the moving window of 20 returns.
Figure 2: Median price series (in black), median fundamental value series (red), grey area between 1st and 3rd quantiles.
Figure 3: Distribution of deviations from fundamentals.
Figure 4: Distribution of final profit. First tax regime

Figure 5: Distribution of daily tax payment
Figure 6: Cumulative return dynamics during a Bubble, median of returns (red), grey area between 1st and 3rd quantiles.
Figure 7: Typical price dynamic and simultaneous evolution of proportion of HFT as the result of switching mechanism. As our simulator realizes continuous time asynchronous (event-driven) trading, prices are aggregated at half-second time scale in order to improve their readability and comparability. Each point reflects the last fixed market price at a given round. The first two figures correspond to untaxed market. The last two figures correspond to the market under Non-Value-Added Tax.